Listening to Audio-Recorded Speech in a Foreign Language.

An Investigation of Playback Behaviours as Overt Strategies.

María Fernández-Toro

— Volume 1 of 2 —

Thesis submitted for the award of the degree of Doctor in Philosophy

University of Newcastle upon Tyne, Language Centre, May 2001.

0. Information to the reader

0.1 Abstract

Most studies of learner strategies rely on introspection as the principal method of data collection, particularly in the case of listening strategies, given the covert nature of listening processes. This study postulates that the operations performed by listeners on their machines while listening to recorded speech (stop, rewind, play) are overt behaviours that can be related to specific processes/strategies. Methods of investigation based on the notion of "playback listening" (PL) are proposed and evaluated through a set of three experiments.

Subjects were undergraduate students registered on Spanish beginner courses in two British universities. Experiment 1 examined the strategies related to pausing in three conditions (no pauses, teacher-controlled, subject-controlled) using both observational and introspective data (N=39). Experiment 2 analysed the playback patterns and recall scores of 16 subjects in a listening test comprising two recordings and allowing free playback. Experiment 3 analysed the playback patterns and conversations between task partners while working in pairs (22 protocols) on different recordings using free playback.

The methods described succeeded in establishing a number of specific relations between playback strategies and listening processes/strategies. Firstly, the pace of progression in a PL task indicates the quality and quantity of processing (controlled and/or automatic) that is taking place. Secondly, subjects had a tendency (which increased with input difficulty) to play the beginning of recordings more than the end. The most repeated segments were also the best recalled, and intensive repetition was shown to increase the frequency of bottom-up decoding. Thirdly, subjects typically approached tasks in a sequential manner, and the order of test questions partly affected playback strategy. Finally, a number of affective responses to input and task were found to occur before cognitive processing began in a PL task. Teaching recommendations are made on the basis of these findings.

0.2 Contents

VOLUME 1

0. INFORMATION TO THE READER **2**

0.1 ABSTRACT 2
0.2 CONTENTS 3
0.3 LIST OF TABLES AND ILLUSTRATIONS 8
0.4 ACKNOWLEDGEMENTS 10
0.5 LIST OF ABBREVIATIONS USED IN THIS THESIS 11

1. BACKGROUND AND RATIONALE **12**

1.1 INTRODUCTION 12
1.2 USES OF AUDIO-RECORDED INPUT (ARI) AND PLAYBACK LISTENING (PL) 14
1.2.1 Playback listening in everyday life *14*
1.2.2 Audio-recorded input in L2 teaching *16*
1.2.2.1 One feature, a succession of media 16
1.2.2.2 The status of ARI in L2 listening methodology 18
1.2.3 ARI and PL in L2 testing *20*
1.2.3.1 Conventional listening tests 21
1.2.3.2 Integrative listening tests 23
1.2.3.3 PL in the testing literature 25
1.2.4 ARI and PL in listening research *25*
1.3 PL AS A RESEARCH TOOL 26
1.3.1 Playback strategies as observable output *26*
1.3.2 Why limit the study of PL to audio when there is video? *27*
1.4 AIMS AND STRUCTURE OF THIS STUDY 28

2. LISTENING IN A FOREIGN LANGUAGE **30**

2.1 THEORIES OF LANGUAGE COMPREHENSION 30
2.1.1 Verbal memory *30*
2.1.2 Knowledge *33*
2.1.2.1 Declarative and procedural knowledge 33
2.1.2.2 Controlled vs. automatic processing 34
2.1.3 Language comprehension *37*
2.1.3.1 Models of language comprehension 37
2.1.3.2 Decoding 40
2.1.3.3 Inferencing 41
2.2 LISTENING IN A FOREIGN LANGUAGE 45
2.2.1 Cognitive processes in L2 comprehension *45*
2.2.1.1 STM capacity constraints 45
2.2.1.2 Decoding spoken words 47
2.2.2 Effects of temporal variables *48*
2.2.3 Factors affecting difficulty in a listening task *50*

3. PLAYBACK LISTENING (PL) **54**

3.1 DEFINING PLAYBACK LISTENING 54
3.1.1 Nature of the listening input *55*
3.1.1.1 Authenticity 55
3.1.1.2 Orality: scripted, semi-scripted or un-scripted? 59
3.1.2 Modes of presentation of recorded input *61*
3.1.2.1 Who has control over playback procedure? 61
3.1.2.2 What is the playback sequence used? 62
3.1.2.3 What cues is the listener given for schema activation? 64
3.1.2.4 What output is required from the listener? 66
3.2 PL AS A SPECIFIC SKILL 68
3.2.1 Adverse features for L2 learners *68*
3.2.1.1 Lack of visual channel 68

3.2.1.2 Quality of the acoustic signal 69
3.2.1.3 Lack of listener-speaker interaction 70
3.2.2 Facilitating features for L2 learners *71*
3.2.2.1 Listening anxiety 71
3.2.2.2 Repetition in PL 73

4. LISTENING STRATEGIES **77**

4.1 DEFINING "STRATEGY" 77
4.2 LEARNER STRATEGIES FOR L2 LISTENING 80
4.2.1 Identifying and classifying listening strategies *80*
4.2.2 Strategy use in playback listening *83*
4.2.3 Effective and less effective listeners *91*
4.3 STRATEGY TRAINING FOR LISTENING SKILLS 97
4.3.1 Can strategy training improve performance? *97*
4.3.2 Approaches to strategy training for listening skills *99*
4.4 METHODS OF INVESTIGATION IN STRATEGY RESEARCH 104
4.4.1 Common methods used in strategy research *104*
4.4.2 Introspective studies: dangers and precautions *106*
4.4.3 Classifying strategies *107*

5. RESEARCH QUESTIONS FOR THIS STUDY **111**

5.1 EXPERIMENT 1: PAUSES 111
5.2 EXPERIMENT 2: INDIVIDUAL MOVES IN FREE PLAYBACK 113
5.2.1 Taxonomy of playback patterns *113*
5.2.2 Repetition and recall rate *113*
5.2.3 Playback patterns: the "a-b-a" hypothesis *114*
5.3 FREE PLAYBACK IN PAIRS 115

6. RESEARCH INSTRUMENTS FOR THE STUDY OF PL **117**

6.1 OPERATIONALISING MEANING-RELATED VARIABLES IN ARI 117
6.1.1 Recall protocols *117*
6.1.2 Idea-units *117*
6.1.3 Measuring redundancy in PL *120*
6.1.4 Idea-maps *121*
6.1.4.1 Segments 121
6.1.4.2 Idea-maps 122
6.2 OPERATIONALISING STRATEGY-RELATED VARIABLES IN PL 125
6.2.1 Move logs *125*
6.2.2 Talk-aloud protocols from paired listening tasks *126*
6.2.2.1 Rationale 126
6.2.2.2 Conventions for protocol analysis 128

7. EXPERIMENT 1 - PAUSES **136**

7.1 PROCEDURE 136
7.1.1 Data collection: *136*
7.2 RESULTS 139
7.2.1 Perceived usefulness of different playback procedures *139*
7.2.2 Students' reported strategies *140*
7.2.3 Observed pausing strategies in group C *142*
7.2.3.1 Comparing observed and self-reported pausing strategies 142
7.2.3.2 Number and distribution of pauses made 144
7.2.3.3 Pause type and strategic trends 145
7.2.4 Effects of pauses on performance *147*
7.3 DISCUSSION 149
7.3.1 Implications of results for the original hypotheses *149*
7.3.2 Subjects' reported strategies *151*
7.3.3 Hypothetical pausing patterns *152*
7.3.4 Discrepancies between self-reported strategies and observational data *111*

8. EXPERIMENT 2 - INDIVIDUAL MOVES IN FREE PLAYBACK 156

8.1 HYPOTHESES TO BE TESTED 156
8.2 PROCEDURE 157
8.2.1 Data collection: *157*
8.2.2 Data processing: *158*
8.3 RESULTS 159
8.3.1 Task sequencing *159*
8.3.2 Use of straight runs at start and end *161*
8.3.3 Playback frequency *163*
8.3.4 Types of playback patterns *164*
8.3.4.1 Local playback patterns: moves, segments and runs 165
8.3.4.2 Global playback patterns 168
8.3.5 Position of move clusters in the move logs for recordings E and F *170*
8.3.5.1 Recording E ("task one") 170
8.3.5.2 Recording F ("task two") 171
8.4 DISCUSSION 174
8.4.1 Task sequencing *174*
8.4.2 Playback frequency *175*
8.4.3 Playback patterns *176*
8.4.3.1 Progression in cycles 176
8.4.3.2 The a-b-a strategy hypothesis 176
8.4.3.3 Use of straight (or fast) runs in the middle of the task 177
8.4.3.4 Progression in slow runs 178
8.4.3.5 Top-heavy patterns 178
8.4.4 Conclusions of Experiment 2 ..... *179*

9. EXPERIMENT 3 - FREE PLAYBACK IN PAIRS 181

9.1 HYPOTHESES TO BE TESTED 181
9.2 PROCEDURE 183
9.2.1 Data collection *183*
9.2.2 Data processing *185*
9.3 RESULTS 186
9.3.1 Move log analysis *186*
9.3.1.1 Pace-related variables 187
9.3.1.2 Repeat rates 189
9.3.1.3 Summary and discussion 191
9.3.2 Protocol analysis *194*
9.3.2.1 Fast progression at the beginning of a task 194
9.3.2.2 Fast progression at the end of a task 201
9.3.2.3 Top-heavy patterns 204
9.3.2.4 Playback sequence in recording F 210
9.3.2.5 Fast progression overall 216
9.3.2.6 Slow progression overall 221
9.3.2.7 The a-b-a pattern 225
9.3.2.8 Playback frequency and recall 228
9.3.2.9 Social-affective strategies 235
9.4 SUMMARY OF EXPERIMENT 3 FINDINGS 246
9.4.1 The first run in playback listening tasks *246*
9.4.2 The last run in playback listening tasks *247*
9.4.3 Top-heavy patterns *247*
9.4.4 Recording F (response cued by questions): *248*
9.4.5 Fast overall progression *248*
9.4.6 Slow overall progression *249*
9.4.7 The a-b-a strategy (fast-slow-fast) *249*
9.4.8 Playback frequency and recall *250*
9.4.9 Social-affective strategies *250*
9.4.10 Global patterns observed in the four recordings *251*

10. CONCLUSIONS 257

10.1 FEATURES OF PL 257

10.2 METHODOLOGICAL ADVANCES FROM PL RESEARCH 259
10.2.1 Idea-maps 259
10.2.2 Move logs 260
10.2.3 Paired listening protocols 261
10.3 L2 LISTENING PROCESSES AND STRATEGIES IN PL 263
10.3.1 Pace of progression and cognitive processing 263
10.3.2 Patterns of variation in playback frequency 264
10.3.3 Sequential approaches to playback strategy 266
10.3.4 Cognitive processing in the most repeated segments 268
10.3.5 Strategic stages of playback procedure 268
10.3.6 Social/affective strategies in playback listening 270
10.4 RECOMMENDATIONS FOR THE TEACHING AND TESTING OF LISTENING SKILLS 273
10.4.1 Pace of progression 273
10.4.2 Playback frequency and global playback patterns 274
10.4.3 Strategic sequence in free playback listening tasks 274
10.4.4 Paired listening with free playback 275
10.4.5 "Adverse" affective reactions 276
10.5 SUMMARY OF CONCLUSIONS 277

11. REFERENCES 279

VOLUME 2

CONTENTS 2

2. TRANSCRIPTS AND TRANSLATIONS OF THE RECORDINGS USED 4

2.1. RECORDING A1 - TRANSCRIPT AND TRANSLATION 4
2.2. RECORDING B1 - TRANSCRIPT AND TRANSLATION 5
2.3. RECORDING C1 - TRANSCRIPT AND TRANSLATION 7
2.4. RECORDING A2 - TRANSCRIPT AND TRANSLATION 9
2.5. RECORDING B2 - TRANSCRIPT AND TRANSLATION 10
2.6. RECORDING C2 - TRANSCRIPT AND TRANSLATION 13
2.7. RECORDING D - TRANSCRIPT AND TRANSLATION 15
2.8. RECORDING E - TRANSCRIPT AND TRANSLATION 16
2.9. RECORDING F - TRANSCRIPT AND TRANSLATION 16

3. COMPREHENSION QUESTIONS FOR RECORDINGS E AND F 17

3.1. QUESTIONS USED WITH RECORDING E (EXPERIMENT 2) 17
3.2. QUESTIONS USED WITH RECORDING F (EXPER. 2 AND 3) 17

4. IDEA-UNIT CHECKLISTS 18

4.1. RECORDING A1 - IDEA-UNIT CHECKLIST 18
4.2. RECORDING B1 - IDEA-UNIT CHECKLIST 19
4.3. RECORDING C1 - IDEA-UNIT CHECKLIST 21
4.4. RECORDING A2 - IDEA-UNIT CHECKLIST 24
4.5. RECORDING B2 - IDEA-UNIT CHECKLIST 25
4.6. RECORDING C2 - IDEA-UNIT CHECKLIST 28
4.7. RECORDING E - IDEA-UNIT CHECKLIST 30
4.8. RECORDING F - IDEA-UNIT CHECKLIST 31

5. EXPERIMENT 1 - PAUSES: STUDENT QUESTIONNAIRES 32

5.1. GROUP A - STUDENT QUESTIONNAIRE 32
5.2. GROUP B - STUDENT QUESTIONNAIRE 33
5.3. GROUP C - STUDENT QUESTIONNAIRE 34

6. EXPERIMENT 2 - INDIVIDUAL MOVES IN FREE PLAYBACK 35

6.1. USE OF STRAIGHT RUN AS THE FIRST / LAST MOVE OF THE TASK IN RECORDINGS E AND F 35
6.2. PLAYBACK FREQUENCY IN RECORDINGS E AND F 36

7. EXPERIMENT 3 - MOVE LOGS 37

7.1. MOVE LOGS FOR RECORDING D (#1 TO #4) 37
7.2. MOVE LOGS FOR RECORDING C1 (#5 TO #10) 39
7.3. MOVE LOGS FOR RECORDING B1 (#11 TO #16) 41
7.4. MOVE LOGS FOR RECORDING F (#17 TO #22) 43

8. EXPERIMENT 3 - RECALL RATE TABLES 45

8.1. RR TABLE FOR RECORDING D 45
8.2. RR TABLE FOR RECORDING C1 47
8.3. RR TABLE FOR RECORDING B1 49
8.4. RR TABLE FOR RECORDING F 51

9. EXPERIMENT 3 - VERBAL PROTOCOLS 53

9.1. TRANSCRIPT CONVENTIONS (USED IN ALL PROTOCOLS): 53
9.2. ENCODING CONVENTIONS FOR PROTOCOL ANALYSIS 54
9.3. DISPLAY CONVENTIONS FOR PROTOCOL #1 (ENCODED EXAMPLE) 57
9.4. VERBAL PROTOCOLS 58
Protocol no. 1: Alice & Ann (D) 59
Protocol no. 2: Bill & Linda (D) 63
Protocol no. 3: Michael & Paul (D) 71
Protocol no. 4: Chloe & Gill (D) 75
Protocol no. 5: Mark & Liz (C1) 81
Protocol no. 6: Richard & Sam (C1) 93
Protocol no. 7: Michael & Andrea (C1) 101
Protocol no. 8: Delphina & Lesley (C1) 110
Protocol no. 9: George & Neil (C1) 125
Protocol no. 10: Robert (C1) 130
Protocol no. 11: Anna & Will (B1) 135
Protocol no. 12: Yan & Tom (B1) 138
Protocol no. 13: Ron & Richard (B1) 142
Protocol no. 14: Helleni & Ismail (B1) 148
Protocol no. 15: Mary & Kelly (B1) 152
Protocol no. 16: Chris (B1) 157
Protocol no. 17: David & Hugo (F) 164
Protocol no. 18: Neil & Mark (F) 167
Protocol no. 19: Will & Oscar (F) 172
Protocol no. 20: Robert & Sally (F) 177
Protocol no. 21: Chris & Gabriel (F) 186
Protocol no. 22: George (F) 192

0.3 List of tables and illustrations

Table 1-1: Advantages and disadvantages of ARI as an L2 learning resource 19
Table 2-1: Structure of the memory system (adapted from Atkinson and Shiffrin 1968) 31
Table 2-2: Commonly accepted difference between the three stages of verbal memory (Craik and Lockhart 1972:672) 33
Table 2-3: Procedural versus declarative knowledge (adapted from Anderson 1976) 34
Table 2-4: Main features of controlled and automatic processing (adapted from Schmidt 1992) 36
Table 2-5: Text genres and semantic categories associated with text genres (Rost 1990:74) 44
Table 2-6: A framework for grading listening input (adapted from Brown and Yule 1983b:107). 50
Table 2-7: Factors affecting difficulty in L2 listening tasks (based on Anderson and Lynch 1988) 52
Table 3-1: Qualitative effects of different playback procedures (adapted from Sherman 1997) 63
Table 4-1: Learning strategy classification (O'Malley and Chamot 1990, study two: p.126) 82
Table 4-2: Listening strategy classification (O'Malley, Chamot, and Kupper 1989; 1990, study 3) 82
Table 4-3: Levels of subject control over playback (Zhao 1997) 84
Table 4-4: Strategies used by learners of Russian in the Rubin and Thompson's longitudinal study (Rubin and Thompson 1992) 87
Table 4-5: Task-related differences in strategy use (summarised from Defilippis 1980) 89
Table 4-6: Experimental set-up in Merlet and Gaonac'h (1995) 90
Table 4-7: Picture selection strategies (adapted from Ross 1997) 94
Table 6-1: Idea-unit scoring grid 119
Table 6-2: Idea distribution map for recording A2. 124
Table 6-3: An example of move log 126
Table 7-1: Group preferences on playback procedure 139
Table 7-2: A comparison of self-reported and observed pausing strategies in group C 143
Table 7-3: Pauses made by group C 144
Table 7-4: Types of strategic pauses 145
Table 7-5: Proficiency and performance scores of groups A, B and C 148
Table 7-6: The introspecting caterpillar (original) 155
Table 8-1: Task sequencing 160
Table 8-2: Effects of task sequence on general recall 161
Table 8-3: Use of a straight run as the first / last move of the task in recordings E and F (summary of results) 162
Table 8-4: Average playback frequency of segments containing "recalled" and "not recalled" ideas (recordings E + F) 164
Table 8-5: Segment repetition - Recordings E+F 164
Table 8-6: An example of move log 165
Table 8-7: Three runs 165
Table 8-8: A straight run 166
Table 8-9: A run with two breaks 166
Table 8-10: A run with two recaps 167
Table 8-11: A cluster 167
Table 8-12: Progression in cycles 168
Table 8-13: Fast pace strategy 169
Table 8-14: Slow pace strategy 169
Table 8-15: Top-heavy strategy 169
Table 8-16: Sequence of information and related questions in recording F 172
Table 9-1: Pace-related variables for recordings D, B1, C1 and F 188
Table 9-2: Summary of global playback patterns found in recordings D. C1, B1 and F 191
Table 9-3: Grading of difficulty in the four recordings used 192
Table 9-4: Input features in Experiment 3 193
Table 9-5: Use of a fast run at the beginning of a task 201
Table 9-6: Use of a fast run at the end of a task 204
Table 9-7: Most recognised words in recording D 205
Table 9-8: Use of top-heavy patterns 207
Table 9-9: Information sequence in recording F 211
Table 9-10: Question-driven and text-driven sequence features in recording F 212
Table 9-11: Move logs displaying fast overall progression 218

Table 9-12: Move logs displaying slow overall progression *221*
Table 9-13: Decoding tokens in fast and slow overall progressions *223*
Table 9-14: Analysis of most repeated segments *230*
Table 9-15: A comparison of the four recordings used *253*

0.4 Acknowledgements

First of all, I wish to thank my supervisor Philip Shaw, for maintaining his support several years after leaving this university; and Francis Jones for taking over his role for the final arrangements. Thank you as well to Paul Meara for helping me out of the "mid-PhD blues" with his sound advice. I am also indebted to the teachers who allowed me to use their classes for Experiment 1, and particularly to Lola Latorre, who also helped with inter-rater reliability tests. All the students who took part in this study are warmly thanked for their contribution.

This PhD could not have been completed without the very practical assistance of my colleagues at the Language Centre: Ian McArdle, Andrea Wilczynski, Derek Green, Alison Tate and the General Office staff, who made every effort to help me secure precious time for this study.

A very special thank you to my husband Jim Bateman for his patience, advice and support in all those practical, moral, and intellectual aspects of the task over the years.

Finally, many thanks to Patas the caterpillar for his valuable contribution to introspective research.

0.5 List of abbreviations used in this thesis

#14/4	Protocol no. 14, move no.4 (for locating quotes from verbal protocols).
ARI	Audio-recorded input
FL	Foreign language
L1	First language
L2	Second language (this abbreviation is also used for referring to a FL in comparative discussions involving both L1 and the FL).
LTM	Long-term memory (see p.32)
LTS	Long-term store (see p.31)
M	Move, moves (see p.125)
MCQ	Multiple choice question
NL	Naturalistic listening (see p.68)
NNS, NNSs	Non-native speaker, non-native speakers
NRI, NRIs	Non-redundant idea, non-redundant ideas (see p.120)
NS, NSs	Native speaker, native speakers
PL	Playback listening (see p.54)
Qs	Questions (in a listening test)
RI, RIs	Redundant idea, redundant ideas (see p. 120)
s., ss.	Segment, segments (see p.121)
Ss	Subjects (in an experiment)
STM	Short-term memory (see p.32)
ZPD	Zone of proximal development (see p.53)

1. Background and rationale

1.1 Introduction

For a number of decades, recorded speech has been a familar feature for millions of language learners around the World. It is difficult to imagine a programme of instruction in a foreign language that would not use recordings of one kind or another at some point, either as an aid to learning or as a convenient medium for assessment. Successive technological media capable of supporting recorded speech (language laboratories, off-air recordings on audio cassette, video, satellite TV, CD-ROM, downloadable audio and video files) have been appearing over the years. Each of these gave rise to a surge of research interest in the new medium as to its suitability and potential applications for foreign language learning. This is because each one integrated a new, promising feature that the previous medium did not have. Video, for instance, opened a wealth of new applications related to the presence of a visual channel; satellite TV gave on-tap access to authentic material; digital recordings did away with the constraints of sequential access, making interactive resources possible; and now the Internet has put an overwhelming volume of multimedia resources from the entire World at our fingertips. Such accelerated development is understandably exciting for researchers in the field of language teaching and learning. Excitement for the teaching potential of new technologies has a downside however: the focus of interest tends to shift constantly onto the the latest feature ("video", "authentic", "interactive", "multimedia", "on-line") much before the "old" features (which are still present in the newest medium) can be fully understood. One such feature is indeed recorded audio.

Unfortunately the time at which research interest was naturally centered on "the ability to play back" as the most relevant feature of recorded speech coincided with the introduction of language laboratories. In those days, these were used in a very restricted manner, dictated by the audiolingual method that relied on them. The method's typical procedure for playing back a recording allowed no scope whatsoever for applying individual learning strategies. By the time the communicative approach did away with the rigid stimulus-response-feedback routine, a new feature ("authenticity") was already attracting most of the research related to recorded speech. As a result, the processes and strategies that are specific to recorded

speech (i.e. related to the very fact that it can be played back) were never properly investigated. One aim of the present thesis is to fill this existing gap.

There is another reason for this renewed interest in a seemingly "old" medium. The notion of learner autonomy and a recent emphasis on independent and open learning have caused a growing interest in learner strategies. Most of the studies conducted in this area rely very heavily on introspection as the main (or only) method of data collection. This is particularly true of research on listening strategies, given that listening does not naturally produce overt behaviours other than response to the input. Whilst the product of listening comprehension may result in observable behaviour (a point that is often debated), the intervening processes have long been regarded as inaccessible to observation. This thesis postulates that the operations that L2 learners perform on their cassette players while listening to recorded speech (stop, rewind, play) are overt behaviours that can be related to specific listening strategies. The three experiments reported here were designed to explore and evaluate the methodology that developed from this principle. Thus the primary aim of this study is of a methodological nature.

Empirically, the aim of the experiments is to identify the strategies that L2 listeners use when exposed to audio-recorded input that can be played back on a machine, and to examine some of the cognitive processes that are elicited in such mode of exposure. This specific modality of listening will hereafter be refered to as "playback listening" (PL). Finally, the findings obtained are used for drawing up a set of recommendations related to the use of PL in foreign language instruction.

PL cannot take place unless the spoken input has been previously recorded in one way or another, therefore a large part of the present chapter will discuss the role of audio-recorded input (ARI) in our everyday lives, in language teaching, in language testing and in listening research. The reasons for undertaking a dedicated study of PL will be recapitulated after this review, and to conclude the chapter, a set of fundamental questions will be proposed.

The specific features of ARI and PL will be examined in chapter 3.

1.2 Uses of audio-recorded input (ARI) and playback listening (PL)

The use of recorded input is closely tied to the comparatively recent explosion of media technology development. As (Buck 1995:113) pointed out, "it is amazing to note that until the invention of sound recording a few years ago, no human being had ever heard exactly the same text spoken twice". But since Edison demonstrated the first gramophone in 1877, recorded sound has been continuously improving in sound quality, price and portability. Old media may become obsolete as new ones are developed, but in one form or another, recorded audio now reaches the vast majority of the human population and it is here to stay.

Indeed, the main reason argued earlier in favour of studying PL was the wide use that is made of this form of listening, both in L1 everyday life and in the areas of L2 teaching, L2 testing, and listening research. Each of those areas will now be reviewed.

1.2.1 Playback listening in everyday life

This section is not a list of all the possible modalities that audio-recorded input can take in our everyday life. As a matter of fact, a huge section of modern society spends most of its spare time exposed to ARI in the form of pre-recorded television broadcasts (soap operas, feature films, documentaries, etc.). However, one can hardly talk about PL when watching conventional TV, given that the only control the viewer has is to turn the volume up or down, to switch to a dfferent channel, or to turn off the TV set altogether. The list below focuses on a few everyday life situations in which listeners actually have some degree of control over the playback of audio- or video-recorded material.

- *Recordings for domestic entertainment*

The range of programmes available is endless, and the most typical reason for recording a programme, hiring a video or buying it is so that it can be played back at a time that is convenient to the user. In such cases, not much use of playback control is to be expected, except perhaps stopping the tape to answer the telephone or put the kettle on. Occasionally, one may miss out a part of the programme, in which case the tape can be stopped in order to play the relevant sequence again, but this strategy is not likely to be used extensively by native speakers. Another reason for

playing back a section again may be that the hearer/viewer simply finds it particularly enjoyable or interesting, or that s/he wants to draw somebody else's attention to it.

- *Educational recordings*

Leaving aside language learning, the range of audio- and video-based educational material is still considerable. There is no space in this thesis to discuss the issue of PL in other disciplines, but one point is worth noting here: when the material is a one-off broadcast, it is not uncommon for the broadcasting institution to advise learners to record it, so that they can study it at their own pace. The BBC *Learning Zone* and the Open University television broadcasts rely on that principle. Audio cassette recordings are also used for instruction in many disciplines, and the Open University often uses "audiovision" (Rowntree 1986:249), which involves a combination of visual information from printed worksheets and explanations and prompts spoken by a teacher on tape (e.g. "Stop the tape while you read the text").

- *Talking books*

Audio has to be used when the receivers of a text are unable to read it in print form, either because their literacy is not sufficiently developed (e.g. audio-recorded stories for children) or because they are visually impaired (talking books for the blind). The playback strategies of these users should not be very different from the strategies that any other listener would use when listening to ARI for entertainment or educational purposes (see above). Obviously the integration of a written component (such as audiovision worksheets) would not be possible with educational material for these users.

- *Answerphones and automatised telephone transactions*

Recorded messages are spreading through many areas of everyday life. They are typically of a transactional nature (the aim is to pass on information to the caller). In early systems, the caller had little control and could only hear he message again by making another call. Otherwise, the only possible reponse from the listener was either to hang up or leave a spoken message, which would also be of a transactional nature (e.g. *"Please return my call asap"*). These days, tone-dialling enables the caller to choose among a number of options that are listed in the recorded message by pressing a key on their telephone. The option to listen again to the same message is often one of the choices available.

- *Dictaphones and audio for record-keeping*

Many professionals use audio recordings as an alternative to note-taking. To save time, letters can be composed orally on a dictaphone, interviews and meetings can be taped for later reference or record-keeping. They can then be transcribed, translated or summarised at a later stage, either by the participants or by a third party (e.g. secretary).

- *Multimedia and on-line applications*

Recorded audio remains one of the ever-present channels in multimedia, either by itself of coupled with text, still pictures, animation and of course video. One of the most useful developments is the gradual improvement of access to foreign radio through the Internet. Many of the stations provide access to live broadcasts[1] (that can be recorded for later use), and a few offer a selection or pre-recorded items from their programmes. Sound quality on live material is not always ideal, but is constantly improving. Songs are another popular instance of recorded material available on the web. Even without a computer, users are now able to download films and documentaries on digital television for home viewing.

1.2.2 Audio-recorded input in L2 teaching

Given that recorded audio is the most convenient means of bringing a variety of listening input into the classroom, and that ARI and PL allow a high level of control over that input, it is not surprising to find it as a recurrent medium in language teaching. However, the precise status of recorded input within the L2 syllabus and the actual use that is made of playback in each case reflect a wide diversity of approaches.

1.2.2.1 One feature, a succession of media

The early stages of audio recording technology caused an initial interest in its potential for foreign language teaching and testing (Dickinson 1970; Jones 1962; Mueller 1959; Spencer 1964). The first systematic introduction of ARI in the language classroom was a result of language laboratory technology. As Matter puts it, the popularity of language labs "was born from the happy, but shortlived marriage between one particular school of pedagogy, viz. behaviourism, and one particular

[1] See for instance COMFM <http://www.comfm.com> and Kidon Media-Link <http://kidon.com/media-link/>.

linguistic theory, viz. American structuralism" (Matter 1989:112). The names given to the three typical categories of language laboratories are telling:

- *audio-passive (AP)*, meaning that learners simply listen to a recording that is controlled from the main console.
- *audio-active (AA)*, meaning that learners have a microphone to record their responses in the gaps provided while the recording plays on, controlled from the console.
- *audio-active-comparative (AAC)*, meaning that learners have full control over recording and playback operations.

The three underlying assumptions are that listening without giving a verbal response is "passive"; that an "active" learner is one that supplies the correct response to a carefully controlled stimulus within a calculated gap on the tape; and finally that the main reason for using the playback facility is in order to "compare" one's response to the correct form. Surprisingly, the three terms are still used today when reporting the technical specifications of a language laboratory.

Eventually, the popularity of language laboratories declined as a result of growing criticism of the audiolingual method. Since institutions had invested considerable amounts of money in setting up the hardware, the medium continued to be used for listening comprehension exercises and testing. Timid attempts were even made to demonstrate the suitability of language laboratory technology for the communicative approach (Ely 1984), but Ely's interesting suggestions came too late. Whilst language laboratories were to be used in schools for many years to come still, very little mention of them was to be found in the newer literature, except perhaps in the form of warnings against fashionable technologies that fail to deliver on the pedagogical front.

Video gradually replaced them as the focus of interest, and became the new favourite source of recorded input in the language classroom throughout the eighties and early nineties (Allan 1985; Lancien 1986; Lonergan 1982; 1983; 1984; McGovern 1984; Stempleski and Tomalin 1990; Strange and Strange 1991).

Two new features now attracted the attention of teachers and researchers. The first was directly related to the nature of video as a medium: it consisted of the ability to convey non-verbal information through the visual channel (Baltova 1994; Kellerman 1990; 1992; Parry 1984; Rubin 1990; Rubin 1995).

The second focus of interest in relation to recorded input (both audio and video) was a result of the communicative approach, and its emphasis on authenticity. The value of authentic listening input was endlessly discussed in the literature (Bacon 1992a; Bacon 1992b; Baddock 1991; Clarke 1989; Denis 1994; Herron and Seay 1991; Klein-Braley 1985; Long 1991; Peterson 1985; Porter and Roberts 1981; Rings 1986, and the list could be much longer). Such interest in authentic recordings was further encouraged in the nineties, as off-air materials from satellite TV became accessible to a growing number of teaching institutions (Cambier and Wright 1995; Fisher, Lynch, and Allen 1995; Hill 1991; Oxford et al. 1993; Rose 1995).

Meanwhile, computer technology had begun to explore interactive multimedia. The first examples of recorded speech in interactive applications used a cumbersome mixture of computer-controlled audio and video tapes (Janssen and van Loon 1991; Meinhof and Bergman 1991a; Rushoff and Wolff 1991). Such reliance on sequential access caused considerable delays, undermining the freedom of movement through the material that users were supposed to enjoy. It took only a few more years to develop systems that were efficient enough to support user-friendly multimedia. The introduction of digital audio and video (coupled with file compression and larger memory size) did to tapes what word-processing had done to mechanical typewriters a decade earlier: for the first time the constraints of sequential access had been eliminated, and immediate random access to any section was at last possible. At this point, recorded speech effectively entered a new era in which "interactive" became the feature that researchers perceived as most relevant.

The current decade appears to be part of the *dot.com* era, in which the focus of attention is shifting from (stand-alone) "interactive" to the ability to exchange messages, information and resources with any other location in the World.

Throughout the successive periods outlined in this review, a number of features have successively become the main focus of attention: "recorded", "authentic", "visual", "satellite", "interactive", "on-line"... and probably others yet to come. The (deceivingly) humble purpose of this thesis is to return to the oldest of these features (i.e. "recorded") and reassess it in the light of modern developments in strategy research.

1.2.2.2 The status of ARI in L2 listening methodology

Academic standpoints about ARI in the listening methodology literature are varied. They range from Krashen's enthusiastic recommendation for systematic listening

practice with a small number of short recordings (Krashen 1996) to Ur's cautious advice that recordings should only be used "for definite specific purposes", either in cases where it would be difficult or impossible to use live input in the classroom, or for testing purposes.

Table 1-1 below (based on Anderson and Lynch 1988; Rixon 1986; Turner 1995; Underwood 1989; and Ur 1984) summarises the advantages and disadvantages that have been most frequently argued with regard to the use of ARI as an L2 learning resource.

Advantages	Disdvantages
Lack of visual channel (in audio):	
• Better concentration when the speaker cannot be seen, no visual distractions	• Processing is more difficult because visual clues are not available.
In repetition, every exposure is identical:	
• More opportunities for understanding from identical repetition	• The speaker is not able to simplify or rephrase a message if the listener cannot understand it in the first place
Preparation:	
• Teachers can prepare supporting materials that are specific to the difficulties presented	• Teachers have to find/produce/edit suitable recordings, which are not always easy to find.
Quality of the input:	
• Variety of native accents, particularly valuable when the teacher is not a native speaker.	• Sound quality is not always good, especially when using loudspeakers instead of headphones.
Strategies available:	
• Learners can work at their own pace by controlling playback themselves.	• Communication strategies are impossible since learners cannot interact with the speaker.

Table 1-1: Advantages and disadvantages of ARI as an L2 learning resource

The absence of visual cues in ARI is often seen as a reason for preferring video to audio material (Rubin 1995), but some authors recommend that both audio and video should be used, because pictures can be distracting and audio requires greater attention to verbal cues (Rixon 1986; Underwood 1989). The presence/absence of visual cues is also proposed as a feature that can be controlled for grading purposes (Brown and Yule 1983b).

Opinions regarding authenticity are divided, although the general view is that "realistic" input, possibly combining scripted, semi-scripted and unscripted speech is best in the early stages, moving on to more authentic material as proficiency develops (Rixon 1986; Underwood 1989). However certain authors prefer to use mostly live input from the teacher (Buck 1995; Ur 1984). The same authors generally regard repeated playback as unnatural, since identical repetition does not occur in real-life interactive listening. Buck proposes two different uses of ARI, based on Littlewood's (1981) distinction between the phases of *precommunicative practice* and *communicative practice*. The first stage aims "to develop facility with fast natural speech and [...] good listening strategies ". The author claims that "usually it does not take too much practice to learn to do this" (Buck 1995:123 [sic]). The precise status of PL is somewhat unclear, but it would seem that repetition at the precommunicative stage is 'tolerated' rather than encouraged. Once the stage of communicative practice is reached, the position becomes totally clear: "It goes without saying that the texts must be real spoken language [i.e. not scripted?], and that listeners will only hear them once" (p.125).

Other authors have a more favourable attitude to PL. Rixon (1986) admits that ARI does not develop the type of communicative strategies that would be used in real-life interaction, but she also points out the confidence-building effects that repetition may have on learners, especially those suffering from listening anxiety. Frequent mentions of independent learning activities in listening resource books (Rixon 1986; Rubin and Thompson 1994; Turner 1995; Underwood 1989) indicate that playback under the learner's control is considered beneficial.

The common denominator in this brief review is that, whatever their position with regard to the status of PL, all the authors cited above assume without exception (albeit reluctantly) that recordings are used in listening instruction at one point or another. No formal instruction in listening can operate without the variety and flexibility that the use of recorded input provides.

1.2.3 ARI and PL in L2 testing

The advantages of using recorded input for testing purposes were seen as soon as the technology became available in language classrooms (Jones 1962). As Thompson (1995:47) points out, if a listening test is to be given to more than one group, the only way to ensure identical test conditions across all groups is to use

recorded input. The use of ARI is therefore common practice in standard listening tests. All listening tests based on this country's National Curriculum use recordings. This type of tests used at school will be discussed below under the heading "Conventional listening tests".

However, the use of recordings for testing listening skills also has detractors, who argue that most listening in real-life is interactive, and that a test based on recorded input is unnatural and fails to test the relevant skills. In line with this argument, one would expect that examining bodies which are specifically concerned with assessing the candidates' ability to use their listening skills in real life situations (e.g. at the workplace) should be reluctant to use what is seen as an unnatural medium in their examinations. This type of tests will be discussed below under the heading "Integrative listening tests".

To illustrate the different positions adopted by test developers regarding the use of recorded input and playback procedure, a few specific examples will now be reviewed. They all belong to standard tests for FL listening skills that are commonly used in Britain for Spanish.

1.2.3.1 Conventional listening tests

- *GCSE*

The most familiar test format for a large number of foreign language learners in Britain is probably GCSE, taken in schools around age 16. A typical listening paper (Oxford Cambridge and RSA Examinations (OCR) 2000c) would consist of nine exercises. The input is carefully scripted and recorded for the purpose, presented in small chunks (usually no more that one sentence, each corresponding to one question item) separated by 12 second pauses. Each exercise comprises between 5 and 10 such items and is presented twice. The chunks become gradually longer through the test, but never exceed an approximate length of 30 or 70 words. All directions are both written and spoken in the target language, except those signposting sections of the paper and announcing playback procedure, which are spoken in English. Obviously, such a test cannot make any claims of authenticity, but the underlying assumption is that learners at this level would find it very difficult to negotiate either longer turns or unscripted speech.

- *A/S and A -Levels*

The same examinations board offers two advanced papers, each comprising separate sections for listening, reading and writing. At the first level (Oxford Cambridge and RSA Examinations (OCR) 2000a), the listening input is largely authentic (e.g. news items), with some purpose-made material (answerphone message at work), and comprises four tasks. Students are given full control of the cassette player, but they have only 90 minutes to complete all sections of the paper. Like in the the GCSE examination papers, responses are kept as short and non-verbal as possible and take a variety of formats (tick boxes, MCQ, short answers to very specific factual questions). No advice is given as to playback strategy. In fact, the specimen paper does not even state that the testees have control of the tape (this information had to be obtained from the OCR website[2] instead).

The second paper (Oxford Cambridge and RSA Examinations (OCR) 2000b) offers better guidance: "The passage lasts just over 3 minutes. You can listen to it as many times as you want, but you should not spend more than one hour to complete this section of the paper"[3]. The total duration is longer that in the first paper (2 hours 45 minutes) and there is only one listening task, but both the input and the responses required are considerably more demanding (e.g. lengthy open-ended replies in L2 to a native speaker's views on the situation of Hispanic immigrants in Los Angeles). The questions appear to match a traditional approach to item design that has been criticised by Brown and Yule: "The questions are typically evenly distributed throughout the text, usually at intervals which correspond to every three or four lines of a written transcript of the tape" (Brown and Yule 1983b:56). The authors' objections to this method lie partly in the fact that such an approach was originally developed for written texts. They also argue that in listening, the point is "to arrive successfully at a reasonable interpretation", and this does not entail trying to work out "*all* that is involved in the literal meaning of the utterance" (ibid., p. 57). The method, however, is still commonly used in many schools and universities, probably because this type of item is easy to develop. Although the practice of basing the questions on written transcripts has been widely criticised, following the sequence of information in the text is generally regarded as good practice (Thompson 1995).

[2] <www.ocr.org.uk>

[3] Original in L2, my translation.

1.2.3.2 Integrative listening tests

In contrast to the conventional approaches described above, other institutions prefer to focus on listening skills within the wider framework of typical tasks faced by listeners in real life, and particularly at work. In these cases, listening skills are assessed as part of integrative tasks such as face-to-face or telephone interaction, or utilising the spoken information for report writing or problem-solving purposes (Schrafnagl and Cameron 1988).

- *Foreign Languages for Industry and Commerce (FLIC)*

The FLIC (Foreign Languages for Industry and Commerce, London Chamber of Commerce and Industry 1999) follows this kind of approach. It offers five levels from ab-initio, up to the equivalent of NVQ level 4. No recordings at all are used for testing listening at the first two levels (Foundation and Level 1): candidates simply respond to "simple, spoken, work related instructions and directions given by the examiner" and take part in a "brief, structured, transactional dialogue" cued by a picture. At these low proficiency levels, emphasis is on pure interactive listening for transactional purposes.

The three upper levels of the same programme do make use of recordings, although playback is never under the candidate's control. Apart from interacting in discussions or role-plays related to business settings, candidates also perform more skill-specific listening tasks based on recorded input. At Level 2, the tape consists of an announcement related to a map provided. They then have to use the information in response to specific requests from the examiner. The tape is played twice. This procedure appears to acknowledge two facts: firstly, the occurrence of non-interactive listening in certain real life situations (e.g. announcements); and secondly, the fact that L2 listeners at this level may not be able to understand that type of message if they only hear it once (Brindley 1998).

At levels 3 and 4, the recording is played only once, and consists of a dialogue representing a problem that might occur in a business setting. Candidates have to take notes and then recall the main points orally, suggesting possible solutions in a discussion with the examiner. The candidate's status as a listener here could be regarded as non-authentic, unless one argues that a similar task may be performed in real life, say by a secretary taking minutes of a meeting. In such case only one exposure would be possible, a feature that is consistent with the playback procedure adopted in the test.

- *Examinations in Languages for International Communication (ELIC)*

Another examining body that favours an integrative approach is the Institute of Linguists. As part of its ELIC programme, the Institute currently offers qualifications at two levels: Intermediate (equivalent to NVQ Level 3+), and Diploma (NVQ Level 4+), that is, the top two levels offered by the London Chamber of Commerce. With only one exception, all modules in which there is a listening component take the form of face-to-face transactional conversation. Module 1 of Intermediate level is the only one that makes use of recorded input, lasting about 5 minutes. The recordings used may take a variety of forms (radio items, taped discussions, interviews, answerphone messages), and the cassette is presented along with one or more written texts in the L2 related to a similar topic. Candidates are given full control of the tape and three hours to synthesise all the information from the sources provided into a single report written in English. In this particular module, listening in itself is not assessed directly. The listening component is only one of the many skills tested (including L1 writing) in a highly integrative task. Assessment is entirely based on the final product of comprehension (i.e. a written synthesis in L1), deriving from a mixture of written and spoken input:

> Successful candidates will demonstrate the ability to extract information, identify ideas and opinions, identify supporting and opposing viewpoints, locate items from more than one source and identify attitudes and conclusions, as appropriate to the source material provided.
>
> (Institute of Linguists Educational Trust 2000:10).

Recorded input is treated here very much in the same way as written input: the mode in which either of these is processed is almost irrelevant, so long as the required information is extracted in one way or another and subsequently utilised. In such a context, effective playback strategies on the part of a candidate could prove essential. One could still argue that free playback of a cassette is not a "real-life" listening task. However, certain professions do require this type of listening skill (e.g. taping interviews for a newspaper article). Such behaviour would make even more sense if the professional is a non-native speaker, since the use of ARI has proved to be an effective compensatory strategy in professional and academic settings.

From this brief review, it can be concluded that even institutions that are very keen to assess interactive listening skills do find a place for recorded input. Although this mode of presentation is confined to very specific instances, it is interesting to see it being used at all. The variety of (often imaginative) approaches to PL that have been developed is also worth noting.

1.2.3.3 PL in the testing literature

Surprisingly, the number of exposures to the input is not a feature that receives much attention in the L2 testing literature. Often the treatment of this feature is comparatively superficial in relation to other test features. One of the 15 questions formulated in "Study two" of Buck's (1990) award-winning thesis is: "How does test method influence the comprehension process?". However no mention is made of playback procedure (e.g. the number of times the input is presented) under the different components determining task specification. Brindley (1998:176) simply notes that "Undertaking a listening test [..] can be very stressful, particularly if only one hearing of a text is allowed and the recorded input cannot be stopped, as is the case with many large-scale high-stakes tests". As Berne suggests, perhaps the time has come to reassess the status of repeated exposure in standardised tests of listening comprehension (Berne 1995:326).

Another relevant question in test design is whether or not to provide listeners with comprehension questions before they listen to the input. The effects of item preview on PL will be discussed in section 3.1.2.3.

1.2.4 ARI and PL in listening research

Thompson's remark about the importance of PL to ensure reliability across all the groups tested would be just as pertinent in the case of tests carried out for research purposes. Only if the input is recorded can researchers claim that subjects tested on different sessions were exposed to identical listening input.

Tapes are also easy to edit and they enable researchers to expose their subjects to carefully controlled input. For instance, they can manipulate variables such as background noise (Hayashi 1991), insert artificial pauses in the recording (Blau 1990, this study), use repeated exposures to identical input (Bransford and Johnson 1972; Hildyard and Olson 1978; Lund 1991a, to name only a few). Speech rate can also be manipulated by means of speech compression devices (Conrad 1989), mechanically (Blau 1990) or digitally (Zhao 1997) slowed-down playback speed.

Even if the input used for a particular experiment does not undergo any significant manipulation, researchers may need to relate the subjects' responses to certain elements of the input. For instance, when measuring the listener's recall of a recorded dialogue, the first step in data analysis may be to draw up a list of idea-units (see Chafe 1985; Crookes 1990; Kroll 1977; Meyer 1975) on the basis of which recall

scores can be calculated. Analysing the input in this way (or in many others) is only possible if the input has been recorded.

In other words, most findings on the listening process - whether they focus on L1 or L2 - are based on results obtained from ARI data. Before such findings can be generalised to real-life situations, the differences between the processes involved in listening to "recorded" versus "non-recorded" input need to be better understood and taken into account. It seems therefore necessary at this stage to investigate in what ways the two modes of listening differ from each other, and how the differences found might affect L2 learner strategies and the language learning process itself.

The implications of using scripted recordings in listening research are a matter that deserves particular attention. These are discussed in in section 3.1.1.2.

1.3 PL as a research tool

The study of playback patterns used by L2 learners when they listen to recorded input can also be used as a research tool to examine L2 listening strategies.

1.3.1 Playback strategies as observable output

One of the problems when investigating the strategies and processes involved in listening is that observational data are far more difficult to obtain than introspective data (Lynch 1998). Like many other researchers, Brown and Yule admit this fundamental hurdle when it comes to listening research: "We may have to come to the rather bleak conclusion that the processes underlying listening comprehension are, in principle, not accessible at all" (Brown and Yule 1983b:145). It must however be noted that a considerable body of reading comprehension research has been successfully conducted by observing subjects' eye movement through written input (see for instance Altmann, Garnham, and Dennis 1992; 1992; Just and Carpenter 1980). Observing the listeners' moves through a recording as they use stop-rewind-play operations is the nearest equivalent in listening to what gaze studies do in reading. PL research too should reveal how a subject's attention shifts from one section of the input to another, and how long it stays focused on each of these. This is not to say that the type of information obtained from the two methods is identical. Gaze movements are automatic and largely unconscious, whereas the decision to press a button is controlled. Gaze duration is usually counted in milliseconds, whereas the time spent playing back a particular segment will be measured in seconds or even minutes. Therefore it seems reasonable to predict that the type of

information that playback behaviour is likely to supply will mostly relate to controlled, rather than automatic processing. It is probably still true that many of the processes underlying listening remain inaccessible to observation, but the picture might not be quite as "bleak" as Brown and Yule's remark suggests.

A number of researchers have examined listening strategies related to the use of recorded input by means of introspective methods (Defilippis 1980; 1991a; Laviosa 1991b; 1993; Vandergrift 1996; 1998). The use of introspection to investigate PL is discussed in section 4.4.2. The present study adopts a different approach to data collection: rather than trying to elicit introspective accounts of (often covert) strategies, it focuses on the overt behaviours displayed by listeners in PL. The instruments used for that purpose will be described in Chapter 6.

1.3.2 Why limit the study of PL to audio when there is video?

The brief history of audio-recorded input outlined above shows that every time a new medium emerges, there is a surge of interest from researchers into its potential for teaching. However, such interest begins to fade as soon as another, more hi-tech medium makes its appearance. This tendency is understandable, and even desirable (as explained in section 1.1), since the teaching potential of new media has to be explored so that our discipline can move forward. Unfortunately, due to the speed of technological progress, this also means that core features of "older" media are abandoned by research before we have even begun to understand the cognitive processes and learning strategies that each medium elicits. And yet, most of the "old" features are still present in the newer media: today's listening input may well be digital, on-line, interactive, multichannel and multimedia, the fact remains that listeners will still be hearing a recorded audio channel most of the time, except when they use one of the few applications that do involve spoken interaction with human interlocutors.

Precisely for this reason, a deliberate choice has been made in the present thesis to study audio rather than video recordings.

There is no doubt that the visual channel offers many important directions for research on L2 learning and acquisition. However, little attention has been given to the fact that using recorded input instead of "live" or "interactive" listening input (see p.54 for precise definitions) could in itself significantly affect the listening processes involved and the strategies learners use.

By leaving out the wide spectrum of additional variables that the visual channel would inevitably generate, we should be in a better position to examine two issues of considerable relevance for L2 listening research:

1. How the use of *recorded* input in itself may affect the listening process,
2. What specific listening strategies it elicits from L2 learners.

In the process of examining these two central questions, another subsidiary issue will need to be addressed:

3. How observational methods from PL can be operationalised for the purpose of listening strategy research.

The research instruments produced in response to this last question are expected to be applicable to other areas of listening research beyond the scope of this thesis.

1.4 Aims and structure of this study

This thesis aims to investigate four questions in which PL could be particularly relevant:

1. What are the specific features of PL?
2. How can PL be used for research on listening processes and strategies?
3. What does initial PL research tell us about L2 learners' listening processes and strategies?
4. How can PL be used in order to develop listening skills in L2?

A number of issues of potential relevance to question 1 are discussed in chapters 1-4. Chapter 3 focuses particularly on the specific features of PL.

Methodological considerations related to question 2 are presented in chapter 6, and demonstrated in the three experiments reported in subsequent chapters.

Question 3 is addressed in chapters 7-9, which report three experiments based on PL. The aim of these experiments is to apply, test and improve the proposed research methods:

- Experiment 1 - Pauses (chapter 7)
- Experiment 2 - Free playback in test conditions (chapter 8)
- Experiment 3 - Free playback in pairs. (chapter 9)

The first two studies are hypothesis-generating, while Experiment 3 tests some of the hypotheses generated in experiments 1 and 2.

Question 4 is discussed in chapter 10, which evaluates the findings of this study and makes a series of recommendations for the use of PL in L2 teaching.

2. Listening in a foreign language

2.1 Theories of language comprehension

Current models of language comprehension are based on computer systems that try to simulate the way in which the human brain works. This enables researchers to test their theories and gradually identify and understand the complex array of inter-related processes taking place. The most commonly used models in current research derive from connectionism (McClelland and Elman 1986; McClelland and Rumelhart 1981; McClelland, Rumelhart, and Hinton 1986). One of the advantages that have been argued for such models is their "biological plausibility", in other words, the processing units that they use operate in a very similar way to nerve cells in the human brain (Garnham 1994).

Neurons (the cells in the nervous system) are in physical contact with hundreds of other neurons around them. They are connected to each other by chemical substances (neurotransmitters), that are capable of transmitting variations in electric potential (nervous impulses) from one cell to another. Such variations cause the connections between neurons to be either *activated* or *inhibited:* the more nervous impulses per second a cell sends to its neighbours, the stronger (more "active") the resulting connection will be. When sufficient potential is accumulated by a given cell, it passes on a signal to other cells, and the process may continue in that way until complex networks of active connections are formed (Anderson 1985). Such networks are organised into information sequences connecting one "node" to another. Some of these sequences contain "learned responses that direct internal processes or overt actions", whereas others are "purely informational" sequences that do not direct any particular responses (Shiffrin and Schneider 1977:158). McClelland and his colleagues also see the network as a series of "units" that may also take a variety of forms, including "possible goals or actions" and "possible hypotheses" (McClelland, Rumelhart, and Hinton 1986:10). Their theory, which is one of the variations of connectionism, is discussed on p.38.

2.1.1 Verbal memory

In order to handle all the processes involved in language comprehension both effectively and efficiently, Atkinson and Shiffrin (1968) proposed that human memory

uses three different "stores". Each of these serves different purposes and operates in a different way to the others:

1. *The sensory store* (sometimes also known as the "sensory buffer", the "sensory register" or "echoic memory" when referring to listening processes).
2. *Short term memory (STM)* or the short term store (STS). The term "working memory" (WM) is sometimes preferred (Greene 1987:39; Harrington and Sawyer 1992)[4].
3. *Long term memory (LTM)* or the long term store (LTS).

A simplified representation of the Atkinson and Shiffrin model is shown below in Table 2-1.

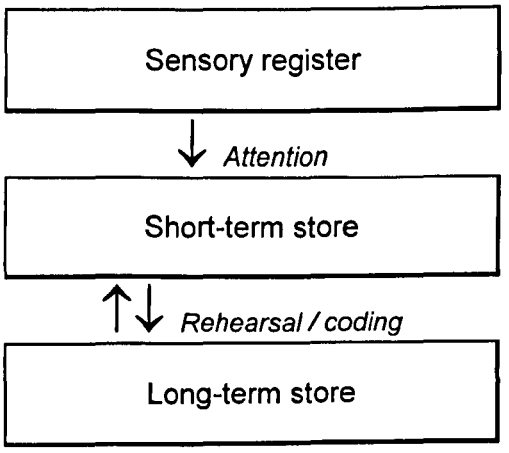


Table 2-1: Structure of the memory system (adapted from Atkinson and Shiffrin 1968)

The sensory buffer receives incoming information in auditory or visual form, and holds it intact until STM can begin to process it. Decay is extremely fast, and information in the sensory buffer can only be held there between 250 milliseconds and 2 seconds. Nevertheless this short period considerably extends the amount of time that a reader or listener has got for identifying information in the incoming signal (Reed 1991:5). In order for information to be transferred into short-term memory, the subject needs to focus his/her attention on it. Any information that is not attended to is immediately lost.

[4] The disctinction between STM and WM is not discussed here as it has little relevance to the present study. See references cited for a discussion.

Short term memory uses incoming information from external stimuli as well as information temporarily retrieved from LTM. Shiffrin and Schneider define it as "the set of currently activated nodes", which acts "as a selective window on LTS to reduce the amount of information for processing to manageable proportions". Unless attention is directed to the active sequence, its period of activation can be as brief as a few milliseconds and the subject will be unaware that it was ever activated (Shiffrin and Schneider 1977:155). If the subject directs his attention to the information, the period of activation in STM may last up to 30 seconds (Craik and Lockhart 1972). But the main constraint affecting STM is its limited capacity. Miller (1956) found that the number of separate items that a subject can hold in STM is between 5 and 9 ("the magical number seven, plus or minus two"), although several items can be grouped together to increase that capacity by means of a process known as "chunking". The existence of a STM separate from LTM has since been contested, but this distinction is of little relevance to this study.

One feature that is of key importance in language comprehension is the fact that working memory has limited capacity, and is forced to perform under considerable pressure. Therefore the processes required in order to comprehend an incoming message have to be prioritised according to the processing resources available at the time (Just and Carpenter 1992). In the case of L2 listening, the constraints will be even heavier, given the lack of sufficient linguistic resources and the pressure of on-line processing. As new information keeps coming in, old information is displaced and lost, unless it has been transfered to LTM before this happens (Call 1985).

Long-term memory comprises all information and responses that have been "learned". Learning is a process of transfer from STM to LTM, that "consists of the association (in a new relationship) of information structures already in LTS" (Shiffrin and Schneider 1977:157). Learning may occur as a result of rehearsal (repetition of information), coding (placing information in the context of other, easily retrievable information) or imaging (creating visual images that help remember) (Atkinson and Shiffrin 1968). Craik and Lockhart (1972) showed that, rather than the number of times a particular item was rehearsed, it was the level of processing that determined retention in LTM: the deeper the processing, the stronger the connection. The capacity of LTM has no known limits, although subjects may be unable to retrieve information in certain cases. This is known as "forgetting".

Table 2-2 (above) summarises some of the features of the sensory store, STM and LTM that have been discussed.

Feature:	**Sensory store**	**Short term memory**	**Long term memory**
Entry of information	Preattentive	– Requires attention	– Rehearsal
Maintenance of information	Not possible	– Continued attention – Rehearsal	– Repetition – Organization
Format of information	Literal copy of input	– Phonemic – Probably visual – Possibly semantic	– Largely semantic – Some auditory and visual
Capacity	Large	– Small	– No known limit
Information loss	Decay	– Displacement – Possibly decay	– Possibly no loss – Loss of accessibility or discriminability by interfer\ence
Trace duration	1/4 - 2 seconds	– Up to 30 seconds	– Minutes to years
Retrieval	Readout	– Probably automatic – Items in consciousness – Temporal/ponemic cues	– Retrieval cues – Possibly search process

Table 2-2: Commonly accepted difference between the three stages of verbal memory (Craik and Lockhart 1972:672)

2.1.2 Knowledge

At this point one needs to consider what it is exactly that the brain stores, retrieves and produces in these different types of memory. This section will examine two important notions: the distinction between declarative and procedural knowledge; and the existence of two types of processing: automatic and controlled.

2.1.2.1 Declarative and procedural knowledge

Knowledge can be divided in two categories: declarative and procedural (Anderson 1976; 2000)

Declarative knowledge	Procedural knowledge
• is represented in terms of a propositional network: ***Propositions*** are abstract, have truth values and have rules of formation.	• is represented in terms of productions: ***Productions*** are set up to encode actions that have proved successful in response to the data available.
• is possessed in an "all-or-none" manner.	• can be partially possessed.
• is acquired suddenly by "being told".	• is acquired gradually by performing the skill.
• can be communicated verbally.	• cannot be communicated verbally.

Table 2-3: Procedural versus declarative knowledge (adapted from Anderson 1976)

Declarative knowledge is explicit knowledge about facts and things, that we can report and of which we are consciously aware; whereas procedural knowledge is knowledge of how to do things. In other words, "knowing what" versus "knowing how" (cf. Greek *episteme* vs. *techne*). Using an analogy from computer systems, declarative knowledge would correspond to data, whereas procedural knowledge would correspond to the programme that uses such data. The summary shown in Table 2-3, which is based on Anderson's ACT theory (Anderson 1976) compares the features of declarative and procedural knowledge.

To take a simple example from musical knowledge, consider a young musician learning the C minor scale on a saxophone. After reading the relevant directions in a teach-yourself method, this person should be able to explain which fingers are supposed to press which keys in what sequence (declarative knowledge). However, it will take many hours of practice until the same person can actually perform the actions described in the book and play the scale at the desired speed (procedural knowledge). The intervening process, known as proceduralisation, is related to the notion of controlled versus automatic processing. These two modes of processing are examined in the next section.

2.1.2.2 Controlled vs. automatic processing

In 1977, Schneider and Shiffrin published the results of a series of experiments in which subjects were required to search or detect certain "target" characters that were

presented among other "distractor" characters on a screen (Schneider and Shiffrin 1977; Shiffrin and Schneider 1977). Their findings confirmed the existence of two "qualitatively different mechanisms", which they called *controlled processing* and *automatic processing* and defined as follows:

> A controlled process is a temporary sequence of nodes activated under control of, and through attention by, the subject. Because active attention by the subject is required, only one such sequence at a time may be controlled without interference, unless two sequences each require such a slow sequence of activation that they can be serially interwoven. Controlled processes are therefore tightly capacity-limited, but the cost of capacity limitations is balanced by the benefits deriving from the ease with which such processes may be set up, altered and applied in novel situations for which automatic sequences have never been learned. Controlled processing operations utilize short-term store, so that the nature of their limitations is determined at least in part by the capacity limitations of short-term store (Schneider and Shiffrin 1977).
>
> An automatic process can be defined [...] as the activation of a sequence of nodes with the following properties: (a) The sequence of nodes (barely) always becomes active in response to a particular input configuration, where the inputs may be externally or internally generated and include the general situational context. (b) The sequence is activated automatically without the necessity of active control or attention by the subject. (Schneider and Shiffrin 1977:2)

The authors also point out that "not all control processes are available to conscious perception". They distinguish two further categories within controlled processing: "accessible" controlled processes (usually slow, easily modified through instruction, e.g. rote learning), and "veiled" controlled processes (too fast to be perceived through introspection and difficult to modify, e.g. serial comparison of items in STM). (Shiffrin and Schneider 1977:159).

In order to perform a processing operation with fluency (Schmidt 1992), a learner needs to perform the same type of processing repeatedly, until it becomes gradually automatised through practice. Automatisation implies that the subject no longer needs to call upon propositional knowledge in order to perform a task, because all the necessary steps are stored in the form of productions. In his ACT* theory of skill acquisition, Anderson (1982; 1989) postulates the existence of two intervening stages for a controlled process to become automatic. The first stage, which he calls *knowledge compilation* involves two mechanisms: series of successive steps in the procedure are collapsed into single productions (*composition*), and factual knowledge itself becomes embedded into productions (*proceduralisation*). The second stage, in which the use of existing productions is adjusted for optimum efficiency, is called *tuning*.

The following features have been identified as characteristic of controlled and automatic processing:

Controlled processing is...	Automatic processing is...
• slow and inefficient	• fast and efficient
• effortful	• effortless
• limited by the capacity of STM	• not limited by short-term memory capacity
• largely under subject control	• not under voluntary control
• flexible	• difficult to modify or inhibit
• at least partly accessible to introspection	• unavailable to introspection

Table 2-4: Main features of controlled and automatic processing (adapted from Schmidt 1992)

These features could be accounted for in two ways:

1. Automatic processing is faster because the rules are applied more efficiently. This is made possible through a process of strengthening connections as a result of repeated exposure (Shiffrin and Schneider 1977). It is to this process that the term automatisation generally refers.
2. Automatic processing is faster because the rules do not need to be applied each time in order to achieve a result. The output of previous operations is directly retrieved from memory. Applying a series of rules (e.g. calculating 7 x 3 by adding 7+7+7) is not as fast as retrieving ready-made information without any further processing (accessing the value 21 directly on a memorised times table).

The two strategies are not mutually exclusive, and each may occur through the different stages of skill acquisition. A shift from one form of processing to another, more effective one, is known as restructuring (Cheng 1985; McLaughin 1990). Performance may be temporarily slowed down as a result of restructuring, because newly-introduced strategies require some practice before they can produce significant effects in terms of improved efficiency (McLaughin 1990; Schmidt 1992).

In a review that focused on the process of speech production, Schmidt (1992) examined the features that constitute fluency. Basing his argument on those aspects of fluency that cannot be directly related to automaticity alone, he pointed out that

"virtually all complex tasks require a mixture of automatic and controlled processes" (Schmidt 1992:362).

2.1.3 Language comprehension

This section discusses the processes that take place in listening comprehension. It begins with a summary of relevant models that have been developed to explain how these processes operate: Levels of Processing theory (Craik and Lockhart 1972); connectionist models (McClelland and Elman 1986; McClelland, Rumelhart, and Hinton 1986); and Anderson's three-stage model of comprehension (Anderson 1985)[5]. Two main types of processes are examined successively: decoding processes (which primarily operate bottom-up), and inferencing processes (which primarily operate top-down). The section concludes with a discussion of the specific constraints that affect these processes in the case of L2 listening.

2.1.3.1 Models of language comprehension

2.1.3.1.1 Levels of processing

The multi-level description of memory presented in section 2.1.1 was criticised by Craik and Lockhart (1972), who regard the division into sensory, short-term and long-term stores as somewhat simplistic. Instead, they prefer to view memory "as a continuum from the transient products of sensory analyses to the highly durable products of semantic-associative operations". The key notion of their theory is the existence of a hyerarchical series of *levels of processing*, in which "greater 'depth' implies a greater degree of semantic or cognitive analysis.(Craik and Lockhart 1972). At the surface would be sensory perception and perceptual analysis (e.g. hearing sounds and recognising phonemes); at the deepest levels would be the global comprehension of a full message, and deeper still the body of world knowledge to which the new information relates. Craik and Lockhart postulate that in order to acquire new knowledge, one has to engage in deep levels of processing. Rehearsal alone does not lead to learning, it simply delays decay. An item will only be learned if deeper analysis of the stimulus takes place before decay.

The existence of different levels of processing is also the underlying assumption in a commonly used metaphor which describes comprehension in terms of *top-down* and

[5] Models of word recognition such as the logogen model (Morton 1969; Morton 1979; Morton 1982), the cohort model (1980; Marslen-Wilson and Welsh 1978) and the autonomous search model (Forster 1976; 1979) are not discussed here for reasons of space. See Pisoni & Luce (1987) for a review.

bottom-up processes. However, in this case the lowest levels (*bottom*) are those levels that are closest to sensory perception (hearing the sounds), whereas the highest (*top*) are levels in which a global meaning is constructed (understanding a whole message, relating it to what we already know). Therefore bottom-up processes are processes that originate in the input, which is decoded into a succession of meaningful units; whereas top-down processes originate in the subject's current knowledge, which is applied to incoming information, in order to interpret the meaning of the message. Although the top-down\bottom-up distinction is only a metaphor and cannot be regarded as a theory of cognitive processing, it is widely used in the literature (Carrell, Devine, and Eskey 1988; Grabe 1991; Rost 1990).

2.1.3.1.2 Connectionist models

Closely based on what we know about the physiology of the human nervous system, connectionism is "concerned with ways of connecting neural elements together to acount for high-level cognition" (Anderson 1985).

McClellend, Rumelhart, and Hinton (1986) developed one of the existing frameworks for connectionist models, called Parallel Distributed Processing (PDP). These models explain information processing in terms of a series of interactions between large numbers of simple processing "units". Each unit sends signals to other units, which may either excite or inhibit the existing connections.

The nature of such units can vary, as the authors explain:

1. In some cases, the units stand for *possible hypotheses* about such things as the letters in a particular display or the syntactic roles of the words in a particular sentence. In these cases, the activations stand roughly for the strengths associated with different hypotheses, and the interconnections among the units stand for the constraints the system knows to exist between the hypotheses.
2. In other cases, the units stand for *possible goals and actions*, such as the goal of typing a particular letter, or the action of moving the left index finger, and the connections relate goals to subgoals, subgoals to actions, and actions to muscle movements.
3. In still other cases, units stand not for particular hypotheses or goals, but for *aspects of these things*. Thus a hypothesis about the identity of a word, for example, is itself distributed in the activations of a large number of units."

(McClelland, Rumelhart, and Hinton 1986:10, my layout and emphasis).

In other words, processing units may stand for any item related to declarative knowledge ("hypotheses") or procedural knowledge ("goals" and "plans"), or for any of the individual features that make up a given item. Connectionist models have been

found to be quite successful in accounting for processes that take place at word level, but less so for syntactic and message-level processing (Garnham 1994).

Focusing specifically on word recognition, Anderson reports an experiment in which Warren (1970, cited in Anderson 2000:65) presented subjects with sets of sentences in which one phoneme (shown below as "•") had been replaced by nonspeech, for instance:

1. It was found that the • eel was on the axle.
2. It was found that the • eel was on the shoe.
3. It was found that the • eel was on the orange.
4. It was found that the • eel was on the table.

Subjects reported hearing "wheel", "heel", "peel" and "meal" respectively. This implies that the word in question was not identified instantaneously on the basis of categorical perception alone (since only two out of the three phonemes required were audible), but that the next lexical word had been decoded before full word recognition could be achieved, thus providing a context that made it possible to infer the identity of the missing phoneme. In other words, processing the segment "• eel" involves two different processes (top-down decoding and bottom-up inferencing), and these operate in parallel rather than in strict temporal sequence. McClelland and his colleagues (McClelland, Rumelhart, and Hinton 1986:12) do admit that processing is sequential to some extent (i.e. there are "transitions from state to state" in the network of currently activated nodes). However, they argue that, even to perform the simplest cognitive tasks, the processes required are so many, so complex and so inter-dependent, that they simply could not operate in a serial manner and achieve the same results.

2.1.3.1.3 Anderson's three stage model of comprehension

Based on connectionism and parallel processing theories, Anderson (1985) analyses the process of comprehension into three stages: perceptual processing, parsing and utilisation.

1. *Perceptual processing*, is represented in the form of the processes "by which the acoustic or written message is originally encoded"[6].

[6] The term "encoding" is often ambiguously used in the literature (as is the case in this quotation). In the rest of this thesis, the term "encoding" will be reserved for the processes that operate in speech production, whereas "decoding" will be used to refer to the processes that listeners perform.

2. *Parsing* is the process by which "the words in the message are transformed into a mental representation of the combined meaning of the words".
3. *Utilisation* is the process "in which comprehenders actually use the mental representation of the sentence's meaning" (i.e. obey an order, answer a question, infer speaker's mood from a statement, take in new information, etc.).

These three stages are said to be partly sequential and partly overlapping in time (Anderson 1985:379).

The next two sections will discuss two main groups of processes that are known to take place in close interaction during listening comprehension: decoding and inferencing. Although both of them can operate at different levels, decoding is mostly related to bottom-up processes, whereas inferencing is mostly related to top-down processes.

2.1.3.2 Decoding

The first step in processing an auditory signal is to recognise the sounds that make up words in the relevant language. This is known as categorical perception. Infants are sensitive to virtually any change in the speech sounds presented to them (whether or not these sounds are used in what is to become their L1). Adult speakers on the contrary always perceive a phonetic stimulus as a member of a particular category. They can easily discriminate between the different phonemes used in their L1, but not among different sounds that would belong to the same category in the L1 (Kess 1992). For instance, Japanese speakers find it difficult to hear the difference between the words "rice" and "lice", because in Japanese the sounds [r] and [l] belong to the same category. This phenomenon, known as categorical perception, is an essential condition for speech perception. The shift towards L1-specific categorical perception occurs around the age of 10-12 months (Vihman 1997). Categorical perception is necessary because no two speakers pronounce the same phoneme in exactly the same way. Furthermore, as we combine different phonemes to make up a string of connected speech, each individual phoneme will undergo a series of changes while it is articulated. Such changes are determined by the movements that the speaker needs to make in order to articulate the neighbouring sounds. For instance, when analysing a spectrogram of the word "bring", it is difficult to tell where exactly [b] ends and [r] begins, since an instant shift from one articulatory position to the next would be physically impossible (Rost 1990). The gradual shift from one to the other is known as co-articulation.

Once the sounds have been recognised, the string of speech has to be segmented (broken up into separate words). Norris and Cutler (1995) produced a computational model of segmentation in two stages. In the initial stage a set of possible "candidate" words are selected on the basis of an acoustic-phonetic analysis of the input. After that, knowledge of the typical metric patterns of the language is used in order to find which of the candidates best matches a typical metric pattern. Norris and Cutler stress that their computational model operates in an entirely bottom-up fashion. Other authors argue that, even this close to the "bottom", some amount of top-down processing does take place as well, and that context is used in order to increase activation for some candidates and lower it for some others (for a full discussion, see Tyler and Frauenfelder 1987)

The outcome of successful segmentation is *word recognition*, which is defined by Tyler and Frauenfelder as "the end point of the selection phase when a listener has determined which lexical entry was actually heard" (Tyler and Frauenfelder 1987).

The next stage is *lexical access*, "the point at which the various properties of stored lexical representations - phonological, syntactic, semantic, pragmatic - become available" (ibid.). For example, when one hears "The cat sat on the mat", word recognition occurs when "mat" is selected over other candidates such as "mattress", "automatic" or "matter". Lexical access occurs when the listener imagines a mat and activates all his/her existing knowledge about the word "mat" (noun, singular, object, a cat may sit on it, rhymes with "flat", etc.).

Once individual words have been decoded, groups of words can be processed and transformed into a mental representation of their combined meanings. This is the process that Anderson defines as *parsing* (Anderson 1985). For instance, "on"+"the"+"mat" is transformed into the single concept of a specific location situated on top of the mat in question.

2.1.3.3 Inferencing

The processes described so far seem to imply that listening is merely a matter of successively combining decoded units into larger units (sounds > words > sentences > text), until the meaning of the original message has been entirely decoded. If listening operated exclusively in such a bottom-up manner, comprehension would be impossible, because much of the meaning of a message is not stated explicitly in the words spoken, but implied by the speaker instead. It is now generally accepted that understanding is not about trying to work out the entire literal meaning of an

utterance. Instead, the process is regarded as successful if the listener arrives at "a reasonable *interpretation* of what the speaker intended to communicate" (Brown and Yule 1983b). In order to achieve this, listeners have to fill in all the missing information by making inferences based on what they already know, i.e. the prior text, prior thoughts, or word knowledge (Trabasso and Magliano 1996; Zwaan and Brown 1996).

Rost (1990) distinguishes five types of inferences that listeners make:

1. estimating the sense of lexical references;
2. constructing propositional meaning through supplying case-relational links;
3. assigning a 'base (conceptual) meaning' to the discourse;
4. supplying underlying links in the discourse;
5. assuming a plausible intention for the speaker's utterances;

Types 1 and 2 above help listeners interpret "the propositional sense of what the speaker says"; types 3 and 4 are employed "to impose a base meaning"; and type 5 assigns the message its "interpersonal relevance". (Rost 1990:63). When encoding inferencing strategies in this thesis, the first two types will be categorised as *local inferences* (propositional meaning) whereas the term *global inferences* (base meaning) will refer to the last three.

In a study based on think-aloud protocols from an L1 reading task by college students, Trabasso and Magliano (1996) identified four types of "thoughts" related to inferencing: paraphrases, associations, explanations and predictions:

1. A *paraphrase* is a transformation of the focal sentence that preserves its gist or meaning.
2. An *association* is an inference that answers questions such as "what, how, where, when and who". It serves to fill in detail on the basis of world knowledge that is activated concurrently to the focal sentence.
3. An *explanation* is an inference that answers a "why" question. It serves to integrate the focal sentence on the basis of previous text information or prior knowlege. It is therefore "backward-oriented" and integrative (creates coherence).
4. A *prediction* is an inference that uses prior text information in order to answer questions regarding the causal consequence of an event, state or action. It is therefore "future-oriented" in relation to the focal sentence. If a prediction is substantiated later in the text, it becomes integrated into the global meaning and creates coherence.

(based on Trabasso and Magliano 1996)

Using the same categories, Zwaan and Brown (1996) found that, when reading short fables in their L1, skilled comprehenders actively generated more explanatory

inferences in order to integrate story events, while less skilled comprehenders generated more associations (a type of inferencing that does not serve an integrative function).

An important condition for inferences to occur is the existence of *schemata*. Schema theory has generated a number of related notions that include *frames* (Minsky 1977), *scripts*, (Schank and Abelson 1977a; 1977b), *scenarios* (Sanford and Garrod 1981), *schemata* (Anderson and Pearson 1988; Bransford and Johnson 1972; Rumelhart 1981). For a further discussion of these notions see Brown and Yule (Brown and Yule 1983a).

The basic principle of schema theory is that all knowledge is packaged into units. When we hear the word "cat", we immediately know (from previous experience) that it is an animal, has two ears and a tail, can jump and climb trees but cannot fly, that its colour may be black, white, ginger, but not green or blue, and so on. Our interlocutor can afford to say "the cat's tail was black" without first having to explain that the cat in question had a tail, because we already know that cats have tails. The same applies to what we know about typical situations: in a "restaurant" situation, we assume a series of roles (customer, waiter, chef) and a series of actions performed in each of the roles (e.g. for the customer: entering, sitting, ordering, eating, paying, leaving). (1977a; Schank and Abelson 1977b).

Rumelhart (1981) lists a series of features that characterise schemata: they have variables (which include a "default" -most typical- value), they can be embedded within each other, and represent knowledge at all levels of abstraction. Schemata are active processes: we use them in order to check current hypotheses and activate or inhibit other schemata accordingly. Rumelhart defines schemata as "recognition devices whose processing is aimed at the evaluation of their goodness of fit to the data being processed". He also points out that schemata may operate in both directions: they may either be "data-driven" (bottom-up), when a sub-schema activates higher level schema; or "conceptually-driven" (top-down); when a schema activates its sub-schemata.

Brown and Yule (1983b) state that native speakers normally encounter spoken language in a specific *context of situation*. As listeners, they apply the stereotypical knowledge that they have being gathering throughout their life about similar situations, and use it in order to construct expectations that help them comprehend what is being said. Listeners use the principles of *analogy* (i.e. assume that things will be the way such things have always been) and *minimal change* (assume that any

differences found in the new situation will be as small as possible). The set of stereotypes related to a given situation is known as its *context*, and comprises schematic knowledge about the speaker, the listener, place, time, genre, topic, and co-text.

Other authors (Zwaan and Brown 1996; Zwaan, Langston, and Graesser 1995) use the term *situation-model*, which is defined as "a coherent mental representation of a narrated sequence of events, actions and states". A situation-model comprises five situational dimensions: time, space, causality, intentionality (related to "motivation"/"goals") and agency (related to "protagonist") - (Zwaan and Brown 1996:312).

A distinction can be made between two types of schemata that we activate in response to written or spoken input: our prior knowledge of the world, known as *content schemata* (Grabe 1991; Long 1989); and our knowledge of the way in which different types of discourse are normally structured, known as *textual schemata* (Long 1989) or *formal schemata* (Carrell 1985; Carrell 1992; Carrell, Devine, and Eskey 1988; Grabe 1991). Rost (1990) explains that listeners anticipate different semantic categories depending on the genre of the input presented (Table 2-5).

The schemata used in narrative fiction have been extensively analysed in *story grammar* research (Black and Wilensky 1979; Mandler 1982; Mandler and Johnson

Genre	*Categories formulated by listener*
• **Descriptions**	Scope of description Identity of individual items Qualities associated with items
• **Narratives**	Event boundaries Characters involved Causes/precedents for actions Outcomes of actions
• **Arguments**	Grounds Claims Warrants Backing
• **Social conversation**	Topic fields Plausible intentions of speaker

Table 2-5: Text genres and semantic categories associated with text genres (Rost 1990:74)

1977; Thorndyke 1977), but as Long (1989) points out, other discourse types of significant relevance to L2 learners have received less attention. Listening research is now also focusing on schemata used in newscasts (1991a; Meinhof and Bergman 1991b; Weissenrieder 1987) as well as academic lectures (Tauroza and Allison 1994). Formal schemata are determining factors in the relative difficulty that a given message presents to the listener (see section 2.2.3 below).

Research findings related to the effects of schema activation on listening performance are discussed in section 3.1.2.3 below.

2.2 Listening in a foreign language

When a subject is listening in a foreign language, a whole set of additional constraints comes into play, causing the processing mechanisms to operate differently from those operating in L1. Typical features of spoken language such as variation phenomena, speech rate and pause phenomena acquire a whole new relevance in the context of an L2. This section concludes with an analysis of the factors contributing to the difficulty of a listening task and a discussion of "level of difficulty" within the framework of Vygotsky's proposed Zone of Proximal Development (ZDP) (Vygotsky 1978).

2.2.1 Cognitive processes in L2 comprehension

The deficits affecting L2 listeners have implications at many levels in the process of listening comprehension. Two aspects of particular relevance will be considered here: problems related to the intrinsic limitations of short-term memory capacity, and problems encountered when decoding spoken words in an L2.

2.2.1.1 STM capacity constraints

As every teacher knows, L2 listeners often complain that the input is "too fast". Processing is more difficult for them because basic low-level processes such as word recognition have not yet been automatised and take them too much time. Given that spoken input arrives continuously at a steady pace, working memory is forced to purge half-processed material before it has been fully decoded, in order to make room for incoming sounds. Measures of STM capacity have been found to be good predictors of L2 proficiency and L2 reading skill (Harrington and Sawyer 1992). Memory for syntactically arranged words (rather than random lists of words, digits or musical tones) was also found to be a good predictor of listening skill (Call 1985).

Just and Carpenter (1992) argue that all individual differences in processing ability are in some way related to STM capacity, and that this theory accounts for individual differences better than any other. According to them, individuals vary in the amount of activation (*total capacity*) that they have available to meet processing demands and also in *processing efficiency* within their available capacity. The authors postulate the existence of "an implicit allocation policy when demands exceed capacity" (for instance in the case of low proficiency in L2). Using a computational model called CC READER largely based on connectionism, they found that in such cases, the processes most likely to be given priority are those that are less demanding on resources (e.g. lower level processes and automatic processes), whereas the higher level processes, which are more resource-consuming are given up (Just and Carpenter 1992:144).

Zwaan and Brown (1996) compared the performance of "skilled" and "less skilled comprehenders" (as measured in terms of their L1 verbal ability) when reading stories in L1 and L2. The aim of their study was to test which of the two following hypotheses about the role of L1 comprehension in L2 reading skill was true:

1. The *linguistic threshold* hypothesis, whereby a minimum level of L2 proficiency has to be attained before readers can achieve comprehension when reading and L2 text (Bernhardt and Kamil 1995).
2. The *linguistic interdependence* hypothesis, whereby L2 reading performance is largely shared with reading ability in L1. Comprehension processes are language independent, not specific to a given language (Horiba, Broek, and Fletcher 1993).

Their results showed that in L1, skilled comprehenders actively generated more "explanatory" inferences in order to integrate story events than did less skilled comprehenders. However in L2, neither group could integrate information across sentences and construct a coherent situation-model. The reason for this is that in L2, both skilled and unskilled comprehenders are severely constrained by a lack of efficient lexical and syntactic processes. This is consistent with capacity theories: in the L2, both groups (skilled and less skilled L1 comprehenders) have to allocate their resources to the generation of an accurate text-based representation. These findings support the linguistic threshold hypothesis. The same study also showed that L1 comprehension skill did have an effect on the construction on L2 bases, which supports the alternative hypothesis of linguistic interdependence.

2.2.1.2 Decoding spoken words

Even at such a low level as auditory perception, the L2 listener is at a disadvantage, given that categorical perception is specific to L1, and develops before we even learn to speak (Vihman 1997). Segmentation too is determined by the specific metrical patterns of our L1 (Cutler 1994; Norris, McQueen, and Cutler 1995), particularly at low levels of L2 proficiency (Delabatie and Bradley 1995).

At word recognition level, the most obvious handicap for an L2 listener is that a variable proportion of the L2 words is simply not known. Kelly (1991) studied the errors made by French native speakers when transcribing English recordings from BBC radio news. The subjects used were moderately to highly proficient in the L2 (38 students having had 5-7 years of instruction in France, and one teacher of English). They were subsequently asked to translate the transcripts that they had produced into their L1. The errors that affected comprehension were found to be mostly related to lexical deficiencies. Perceptual and syntactical errors had less adverse effects on comprehension.

A number of speech features may also prevent L2 listeners from recognising even the words that they are supposed to know. The first category can be summarised under the common term of *connected speech phenomena.* Spoken sounds are produced under certain articulatory constraints (see section 2.1.3.2 above on coarticulation), and this causes neighbouring sounds to influence each other, causing assimilations, elisions, reductions and free variations (Field 1997; Rost 1990:38-9). Another type of change, known as *sandhi-variation,* affects adjacent grammar forms (through contraction, liaison, elision, etc.) thus reducing their saliency to the low proficiency L2 listener (Henrichsen 1984). Variations across different speakers can also be problematic. Some of these (voice quality, articulatory clarity) are specific to individuals, whereas others are shared within a given dialect. Rost (1990) identifies two main factors that can make an accent difficult to understand: the environments in which certain phonemes do or do not occur (*phonotactic distribution*), and *variations in prosodic features* (syllabification, segment duration, pace and stress).

Once a word has been recognised, its meaning needs to be retrieved from LTM. This is a much slower process in L2 than it is in L1, given that connections are weaker in L2 and therefore less easy to activate automatically. When further semantic processing is required (for instance in order to decide whether a target word -"pea"- is a member of a given category -"vegetable"-), subjects can utilise words very effectively in their L1, because once recognised, a word automatically activates its

related nodes in conceptual memory (Dufour and Kroll 1995). However, when the input is presented in L2, low proficiency subjects have not yet established this type of direct link between words in the L2 and conceptual information. Instead, they appear to take a longer route via the L1 to access the conceptual network. That is to say, words are translated in order to gain access to the conceptual network which is available for L1. Dufour and Kroll (*op.cit.*) also claim that some limited access does become possible directly from L2 as proficiency increases. This is because direct connections between the L2 and the conceptual network are gradually being created and strengthened. This shorter route saves processing resources, which can then be used for higher level processes such as inferencing.

2.2.2 Effects of temporal variables

The most characteristic feature of speech in comparison to written language is that it is time-bound. The added constraints that on-line processing imposes on STM capacity have just been discussed. This section reviews research on two temporal variables that are believed to affect listening comprehension: speech rate and pause phenomena. Since many of the studies reviewed involve both variables, they will be discussed under the same heading.

Studies on speech rate and pause phenomena can be found in the L1 literature as early as the sixties (Foulke and Sticht 1969), but L2 research on the same phenomena only began to investigate the matter with suitably accurate methods two decades later (Griffiths 1991). However, as Thompson (1995:39) points out, "research evidence is limited and conflicting because studies use different subjects, languages, texts, tasks, definitions of 'normal' rate for different languages, and measurement techniques".

Common sense intuition would suggest that slowing down speech rate should help L2 listeners. Indeed, Conrad (1989) used speech compression techniques to examine the effect of speech rate on L1 and L2 listening comprehension. Sets of time-compressed sentences were played back at decreasing speech rates and recalls from L1 and L2 subjects were qualitatively and quantitatively analysed at each stage. Not surprisingly, most L1 listeners achieved full recall at much higher compression levels (56% of normal playback time) than L2 listeners, who only reached a maximum of 72% recall even at 91% of normal playback time (the slowest speech rate presented). A similar experiment was conducted by Griffiths (1990) using different

reading speeds instead of speech compression, but none of the differences found reached .05 significance level.

Pauses appear to be more helpful to listeners than slower speech rate. Blau (1990) compared the effects of pause insertion and speech rate on L2 listening comprehension at different levels of proficiency. Speech rate was artificially slowed down by modifying playback speed (using a pitch correction device), whereas pauses were inserted into another recording "at selected sentence, clause and phrase boundaries"[7]. Overall, the pause insertion condition gave significantly higher listening comprehension scores than slower speech rate. Another study by Berquist (1994, quoted in Lynch 1998:9) using similar conditions (normal rate, slowed-down rate, normal rate with pauses) also found performance to be best at normal speech rate with pauses inserted. This is consistent with capacity theories, since a total interruption of input allows time for completing pending processes in STM without competition from incoming stimuli, whereas a continuous stream of speech (even at slow speech rate) does not. Blau also observed a proficiency threshold half-way between high and low levels, at which pauses appeared to be most beneficial for listening comprehension. Speech rate too had different effects at different levels: low-proficiency subjects performed better in the the slowed-down condition, while high proficiency subjects performed better at normal rate (Blau 1990).

It must be noted that pauses made by speakers vary in nature, duration and function. Butterworth (1980) lists three main functions for pauses in speech production: (1) pauses during which words are chosen; (2) pauses during which syntactic plans are formulated; (3) pauses that mark clause endings. The latter type is the only one that is of clear use to listeners. Unfortunately, it is only present consistently in scripted speech. In spontaneous speech, it accounts for no more than one third of all breathing pauses. The remaining pause types (types 1 and 2) only occur in spontaneous speech, where they are also the most frequent types. Although they do give listeners more time for processing, they can also be confusing for NNSs (Voss 1979), since they distort the intonation contour of utterances and may add some irrelevant acoustic information in the case of filled pauses.

Vanderplanck (1993) proposes an alternative to the traditional unit in words-per-minute (wpm) as a measure of speech rate. He argues that listeners are actually

[7] This procedure is very similar to the one used in Experiment 1 of this thesis with group B (p.136).

more sensitive to the *pacing* and *spacing* at which stressed words occur. These variables can be reported as shown in the example below:

- *wpm* = 191 (i.e. 86 words in 27 seconds)
- *pace* = 48 stresses/beats per minute
- *space* = 1:4 (i.e. every 4th word is stressed)

To test this method, he used a group of European students who were asked to shadow a speech by Margaret Thatcher. His findings suggest that what caused time-related difficulties was not the number of wpm in itself. As Vanderplank explains, Mrs Thatcher "has a slow tempo or number of stressed words per minute, but puts in a lot of unstressed words between each stressed word".

2.2.3 Factors affecting difficulty in a listening task

Brown and Yule (1983b) summarised the factors contributing to input difficulty in a framework (Table 2-6 below) comprising two dimensions: level of complexity between different types of input (shown on the horizontal axis), and level of complexity within any one type of input (shown on the vertical axis). The framework was originally designed to represent output difficulty in speaking tasks, but it was later shown to be applicable to the input of listening comprehesion tasks (Brown et al. 1987).

The level of input difficulty of a specific listening task is thought to be a combination of the two dimensions: the top left of the table (i.e. the easiest case) could take the form of one single speaker describing the one simple item, using simple categories (e.g. listing the different rooms in a house); whereas the bottom right corner would

	← Static →		Dynamic	Abstract
	Description	Description / instruction	Storytelling	Opinion-expressing
↑ *less difficult*	**few elements**, properties, relationships, characters or factors which may be difficult to distinguish from each other			
more difficult ↓	**many elements**, properties, relationships, characters or factors which may be difficult to distinguish from each other			
	←	*less difficult*	*more difficult*	→

Table 2-6: A framework for grading listening input (adapted from Brown and Yule 1983b:107).

represent the most difficult case, for instance in the form of a group of speakers arguing the pros and cons of a complex decision. The level of difficulty could be adjusted by modifying the components in one of the two dimensions. For instance, understanding the same group of speakers would be easier if they were discussing a less abstract matter, such as trying to remember the names of people who were at the party yesterday. In the same way, the house description would be harder to understand if the speaker was describing the differences between various bedrooms in the house instead of just listing them.

In a later work, Brown (1994) elaborates on this argument by defining four different types of understanding. These are listed below, graded by increasing difficulty ("all other things being equal"):

1. *Identifying information*[8] - "where a speaker informs a listener of something which requires nothing more than correct identification of the words used" (e.g. phone/car/bank-account numbers, proper names)
2. *Procedural understanding* - "which may only require a minimum of linguistic interpretation" (e.g. directions: how to get to the library; mother to child: "into the bed you go then"). Typically presented one step at a time.
3. *Narrative understanding* - "any genre where the content is temporally organized".
4. *Argument*

The author explains that In the first case, items are merely "identified rather than understood". The next two types are more complex:

> In procedural understanding, you may often close off previous information which records how you arrived at the point where you now are, and no longer retain it in active memory, thus releasing processing resources for the next step. In narrative understanding, however, you must carry relevant previous information in a readily accessible form in memory and, in order to understand fully, constantly relating what is happening now to what happened earlier (Brown 1994:18).

Brown also points out that narratives are easier to understand when events are presented in the order in which they occurred, and more difficult when the sequence is altered by flashbacks and premonitions, because "these require the listener to hold several embedded sub-narratives stably in a mental discourse representation, and to track characters across different time frames". The most difficult of all categories is argument, because...

[8] My term (the author does not supply a specific term for this category).

> Understanding argument requires that you create a mental representation of a number of premises, that you distinguish between these and remember them accurately, and then that you track the abstract relationships established between them, until you reach the conclusion that the speaker wishes you to reach. There is little doubt that abstract arguments, explanations, justifications, theorizing, make more demands on understanding than any other genre. (*ibid.*, p.19]

Anderson and Lynch (1988) integrate Brown and Yule's bi-dimensional analysis into a wider framework comprising not only features of the input itself, but also those related to the nature of the listening task to be performed:

The features of the listening input

- The way in which information is organised
- The listener's familiarity with the topic
- The explicitness of the information given
  - *Redundancy*
  - *Sufficiency of information*
  - *Referring expressions*
- The conceptual complexity of the input (based on Brown and Yule 1983b)
  - *Level of complexity specific to discourse type*
  - *Level of complexity of the input within its category*

Our task or purpose in listening

- Processing load
  - *Amount of information that has to be processed*
  - *Amount of time available*

The context in which listening occurs

- Visual support available
- Group format used (listening individually or with others)
- Type of response required

Table 2-7: Factors affecting difficulty in L2 listening tasks (based on Anderson and Lynch 1988)

There is still another important element to consider in this discussion of listening difficulty: listeners themselves. Proficiency level has been mentioned several times throughout this chapter, but when exactly is a particular recording suitable for a listener's particular proficiency level? According to Krashen's *input hypothesis* (Krashen 1985), in order for language acquisition to take place, the input must contain structures that are only slightly beyond the point that we have currently reached within the "natural order" of language acquisition (a level commonly known as *i+1*). Krashen's notion of *comprehensible input* also postulates that such i+1

structures must appear in a context that makes them comprehensible to the acquirer. The problem with the input hypothesis is that it assumes the *natural order hypothesis,* according to which the rules of a particular language are always acquired in the same order by every individual.

A more flexible approach can be found in Vygotsky's notion of a Zone of Proximal Development, or ZPD (Vygotsky 1978).The theory was developed to account for all areas of child development, and distinguishes at least two developmental levels: *actual development* (what the child is now able to do alone) and *potential development* (what the child is able to do with adequate assistance). Between these two levels lies the *zone of proximal development*, which Vygotsky defines as...

> ...the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers. (Vygotsky 1978:86)

The concept of ZPD has since been applied to L2 learning (Hossein and Cumming 2000), and even to adult L2 learners (de Guerrero and Villamil 2000). Apart from its flexible and learner-centered nature, what makes this theory attractive is that *scaffolding* (assistance that is in tune with the learner's current needs and ability) may be provided not only by an expert, but also by more capable peers (de Guerrero and Villamil 2000) or even - as the present study argues - just peer learners.

3. Playback Listening (PL)

This chapter takes a closer look at the notion of playback listening. First, a basic framework is proposed in terms of the nature of the input used and its possible modes of presentation. The second part of the chapter focuses on the features of PL that may either hinder or facilitate comprehension.

3.1 Defining playback listening

A number of categories have been used in the literature in relation to notions commonly associated with PL. Perhaps the two most common distinctions are *interactive* vs *noninteractive* listening (Rubin and Thompson 1994) and *live* vs. *recorded* (Rixon 1986). However, it must be made clear that neither "recorded" nor "noninteractive" listening convey an identical meaning to PL.

All PL can indeed be characterised as *noninteractive*, insofar as the listener does not have the opportunity to influence the speaker's output (except perhaps with certain computer programmes especially designed to allow a limited degree of interaction, e.g. Zhao 1997). However the reverse is not true: not all noninteractive listening is PL. Listening to a formal lecture is essentially noninteractive, yet it does not constitute PL unless somebody records the lecture in question and plays it back on a machine.

A similar argument applies to *recorded* input: listening to recorded input only constitutes PL if somebody deliberately applies a particular playback procedure in order to control the listener's exposure to such input. A "repeat" TV broadcast or a recorded announcement in a station is indeed recorded material. However, either recording will only be regarded as PL input if a particular playback procedure is deliberately adopted (either by listeners themselves or by somebody else) in order to achieve a specific purpose. The latter scenario (playback controlled by somebody else) is almost exclusively related to teaching and learning situations, whereas listener-controlled PL does sometimes occur in real-life situations involving native speakers (see section 1.2.1).

In the context of L2 learning, the following definitions are proposed:

- ***Playback listening (PL)*** is a type of listening task in which spoken input is presented in recorded form, and where playback equipment is operated in a

particular way (i.e. using"play", "stop", "pause" or "rewind") in order to control the listener's exposure to such input.

- ***Playback listening strategies*** are the strategies that listeners use in order to comprehend the input in a playback listening task and/or to acquire skills or knowledge through a PL task. PL strategies can be applied by listeners whenever they are exposed to recorded input, whether exposure is controlled by themselves or by a third party (e.g. teacher or test administrator).
- ***Playback strategies*** are strategies in which listeners operate playback equipment in a particular way (i.e. using "play", "stop", "pause" or "rewind") for comprehension or learning purposes. Playback strategies can only be applied if listeners themselves have control over playback equipment.

This definition of PL implies the existence of three dimensions in playback listening: the recording used as input, the person controlling the playback procedure, and the playback procedure itself.

Given that any instance of recorded speech is, in theory, susceptible to be used as PL input, section 3.1.1 will focus on two features that have generated considerable discussion in relation to the issues of authenticity and orality. Section 3.1.2 will examine different possible modes of presentation involving teacher- and learner-control over the playback procedure.

3.1.1 Nature of the listening input

3.1.1.1 Authenticity

In order for PL to occur, the input has to be presented in such way as to allow control over the playback procedure. This raises the issue of authenticity: as soon as a message is taken away from the context in which it was originally produced, its authenticity comes into question.

Authenticity may be taken to mean two different things: the first is to do with the original source of the material and its nature, the second with the use made of it. Widdowson (1979; 1984) referred to the first aspect as the *genuineness* of the input, arguing that language that is presented "in its natural communicative condition" needs to be subsequently "authenticated" by the learners within their own frame of reference. Without this step, the original message will not make any sense to them (Widdowson 1984:218). Authenticity therefore is more to do with the relationship that

is established between the input itself and the person using it. Breen (1985) identified four categories of *authenticity*:

1. Authenticity of the texts which we may use as input data for our learners
2. Authenticity of the learners' own interpretations of such texts
3. Authenticity of tasks conductive to language learning
4. Authenticity of the actual social situation in the language classroom.

According to him, the inherent authenticity of the input itself is not as relevant as the degree to which learners may eventually come to interpret its meaning in the same way as fluent language users would. Regarding task authenticity, Breen stresses that meta-communication about learning and about the language ought to occur alongside the task's intended communication output. This becomes particularly relevant provided that the classroom is regarded as an authentic social situation in its own right.

If one analyses PL within Breen's proposed categories, its potential relevance as an authentic learning activity becomes clear:

- *Authenticity of the texts which we may use as input data for our learners*

The input itself will be authentic (in the "genuine" sense as defined by Widdowson 1979) so long as it is recorded in the relevant real-life conditions.

- *Authenticity of the learners' own interpretations of such texts*

Learners will be able to authenticate the input, provided that its content and function are accessible to them. The use of supporting materials and activities may be required in order to achieve this, and PL is one of the methods available. Another teaching strategy consists of using recordings of NSs interacting with NNSs, so that the speakers naturally produce simpler language in the recordings, and learners can identify with the addressee (Hart 1992; Lynch 1989).

- *Authenticity of tasks conductive to language learning*

One could state that a task is authentic when learners are required to assume the same role as the listener would in real-life. On this basis, a first, narrow definition of authenticity could be formulated using MacGregor's proposed categories of listener roles (McGregor 1986) as reference criteria. He suggests that in real-life settings, a listener may act as *participant* (as in a conversation between friends), *addressee* (listening to a formal lecture), *auditor* (listening to announcements in a station),

overhearer (overhearing two strangers talking to each other), *eavesdropper judge* (listening to a recorded conversation in which we may or may not have taken part). So, strictly speaking, only the fifth category would lend itself to authentic listening tasks using PL. However, a more flexible reading of "authenticity" would allow us to include all but those tasks which would require listeners to be "participants" in real time. In the remaining cases, the listener does not influence the speaker's message[9], therefore it would not be unnatural for learners to use recorded input and simulate the role of a real-life non-participant listener.

- *Authenticity of the actual social situation in the language classroom*

The fact that learners assume the role of eavesdropper by listening to taped dialogues is perfectly natural in terms of what Breen calls the "social authenticity of the language classroom". L2 learners are not L1 users. In addition to practice in the target "real-life" skills, they also need opportunities to explore, observe, and discuss the language for themselves. Meta-communication may well be the best means at their disposal for making sense of ("authenticating") the input. PL may not be what a genuine native-speaking listener would do in real life with the same input, but the main issue is really whether or not the use of PL will - as Breen puts it - "help the learner *develop* authentic interpretations" (Breen 1985:63, author's emphasis).

It has already been noted (section 1.2.2.2) that most instructional approaches to listening assume the need to use recorded input in one form or another. The use of PL is an everyday reality in the vast majority of language classrooms. It can therefore be said that for most learners, PL is very much part of their own "real-life" listening experience in the foreign language. To a typical pupil in a typical English secondary school, the real-life experience of French is more likely to be a taped dialogue than face-to-face conversation with a French person. A remark made by Brown and Yule (1983b) implicitly acknowledges the "real" status of PL for FL learners. They point out that recordings of spontaneous conversations such as those used in many listening courses can be very boring for those who are not participants. However, "after a while, anyone who listens to a lot of recorded conversation turns into a conversation analyst and begins to use his analytic observation". They suggest that, if learners are

[9] It is true that even "addressees" can use devices that influence the speaker's output (e.g. a puzzled facial expression), but real-life also has numerous instances in which these are either unavailable (e.g. no eye contact) or ignored by the speaker (e.g. reading out a speech). In other cases, listener responses have to be delayed (e.g. round of questions at the end of a lecture) - (Lynch 1995).

"provided with tasks which help them to become, in however limited a way, conversation analysts", they will learn to find those dull conversations extremely interesting (Brown and Yule 1983b). This is a far cry from "real-life" listening. What it does show is a concern to integrate PL within the learners' situational reality. If PL promotes more explicit analysis than would NL, perhaps listeners should be allowed to pursue this approach in the classroom. Dismissing a potentially useful activity on the grounds that it does not mimic what listeners would do in a hypothetical real life seems somewhat of a luxury, particularly when the learners are not in direct contact with the target culture. Acknowledging PL as a valid mode of listening in formal instruction settings and trying to make the most of it would seem a more realistic option. This of course does not question the desirability of listening tasks of a more communicative nature wherever possible in addition to PL.

Chavez (1998) asked 186 college students of German to rate each of 53 situations (composed of 12 authenticity factors) in terms of their (a) authenticity, (b) contribution to language learning, (c) ease/difficulty, and (d) resulting anxiety/enjoyment. The main factor contributing to authenticity in the eyes of the students was "native inception" (speech produced by a NS). "Native reception" (addressed to NSs) was not so important. No other single variable was a determining factor in itself, but the more factors were present in a situation, the more it was perceived as authentic. "Authentic" generally correlated with "conductive to learning" as well as "enjoyable". Interestingly, there were only two instances in which "authentic" did not mean "enjoyable": the first was related to situations in which learners had to read literary texts without pedagogical support, and the second related to situations involving listening in the absence of visual cues. In other words: watching authentic news on TV was seen as enjoyable, but listening to authentic audio recordings was perceived as anxiety-generating. It must be pointed out that the hypothetical situations presented in the questionnaire did not specify whether or not listeners had control over playback.

In a study involving Hong Kong undergraduates, Lewkowicz (2000) asked the students to identify the attributes of a test that contribute to performance in their view. Authenticity was rarely mentioned as a relevant attribute, although students reported that they would notice its absence if it were missing from a test. Such discrepancy with Chavez' findings (Chavez 1998) could be due to cultural differences between US and Hong Kong students and different teaching traditions in the two countries examined.

3.1.1.2 Orality: scripted, semi-scripted or un-scripted?

Another issue, closely related to the issue of authenticity is the degree to which the language spoken in a recording resembles the language spoken in real life. Mendelsohn (1995:133) complains that "Much of the material is written language that has been recorded in unnaturally enunciated "teacherese" language where, for example, the speech is slower, the pitch movements are greater, and the "citation-form" of the pronunciation used does not contain the fast-speech features of normal spoken English".

Spoken language normally presents higher levels of redundancy (e.g. more repetitions), shorter sentences and fewer subordinate clauses. This reduces the pressure imposed on the listener by on-line processing when memory resources are limited (Just and Carpenter 1992). All of these facilitating devices obviously disappear when the input is read out from a fully scripted text, unless the script was purposefully written with oral output in mind.

Chafe (1985) listed the following features as characteristic of spoken language in comparison to written language:

- Much shorter idea units
- Idea units that are either juxtaposed or co-ordinated (not subordinated)
- Use of flow-monitoring devices (anyway, well now, etc.)
- High occurence of disfluencies due to on-line production
- More colloquial and nonstandard register than written language
- Speakers more personally involved than writers (e.g."I mean"; use of "you", questioning to involve addressee...)
- More use made of categorical truths (direct experience is the main concern), inductive thinking, sensory/hearsay evidence

Tannen (1982; 1985), proposes that, rather than a strictly categorical distinction between oral and written modes, oral texts can be arranged along what she calls "the oral/literate continuum" (Tannen 1980; 1982), in which texts closest to the spoken language would be found at one end, whereas those closest to the written language would be at the opposite end, and comprising a series of intermediate modalities that would be found somewhere between the two extremes. Thompson (1995:37) describes "listener friendly" passages as "those closer to the spoken than the written language" and advises that read-aloud versions of written tests should be avoided as sources of input in listening tests. She also reports higher listening comprehension scores for conversations than for expository passages (Thompson 1993, quoted in Thompson 1995:37).

Based on Tannen's concept of oral/literate continuum, Shohamy and Inbar (1991) compared the listening performance of 150 EFL learners using three different types of input: news broadcast (the most "literate" mode), lecturette (an intermediate mode) and consultative dialogue (the most "oral" mode). They found that performance was significantly better when the information was presented in the form of a lecturette rather than a news broadcast, although no significant differences were found between the lecturette and the dialogue. Their study concluded that texts which are closest to the "oral" end of the continuum are easier for listeners, which confirms the argument against treating orally-delivered written text as another instance of "spoken" text.

A brief comment on methodology seems pertinent at this point. There appears to be one weakness in Shohamy and Inbar's experimental design: the recordings used were fully scripted, read aloud texts:

> Each topic was written in the form of three different text types [...] All passages for each of the topics contained identical factual information and were also controlled for lexis, i.e., the lexical items deemed difficult for the subjects'proficiency level appeared in all three versions of the same topic (e.g., susceptible, phenomenon).- (Shohamy and Inbar 1991:28).

Under such tightly controlled conditions, the speakers used in the experiment would not have been able to display key features of spoken discourse, such as hesitation and pause phenomena directly related to speech production processes. Natural intonation and voice quality would also be lacking, all of which might significantly affect comprehension. Even accomplished actors may find it difficult to mimic these phenomena with total accuracy when delivering a fully scripted text.

A similar comment could be made about an experiment in which Chaudron and Richards (1986) studied the effect of different types of discourse markers on lecture comprehension. In the discussion, the authors argue that "In using micro discourse markers, the speaker is merely succombing to the necessities of on-line discourse production". However such a claim would not really apply in their experimental setting, since all markers had been inserted *a posteriori* into a written script, and subsequently read out in a recording studio.

Studies on the effects of orality on listening comprehension face a methodological dilemma. On the one hand, researchers need to find ways of controlling the vast array of variables related to on-line speech production (hence the common use of scripted input containing the particular features of "orality" under study). On the other hand, as soon as a text is scripted (whether it is read aloud or delivered from memory), it can no longer be legitimately regarded as spoken language.

A similar problem occurs when teachers try to keep the input accessible to the learners by controlling the language used in recorded dialogues through fully scripted text. An alternative that has been proposed (and, I suggest, ought to be adopted for research purposes as well) is the use of semi-scripted recordings. By giving speakers no more than a bare outline of items to cover, one forces them to engage in speech production processes in order to flesh it out. So long as certain basic precautions are taken (Geddes and White 1978), most of the features of spoken language will spontaneously occur, while the desired items (lexical, factual or grammatical) can be controlled by building them into the preset outline.

Finally, it is worth noting that a large proportion of real-life input, such as news broadcasts, screenplays etc. is fully scripted anyway. Good quality instances of these genres ought, however, to be written in a manner that assumes an oral mode of delivery[10] (Gregory 1967:191). It has been suggested that unscripted input is more suitable with advanced levels of proficiency, whereas low-proficiency learners benefit more from scripted ot semi-scripted material (Dunkel 1986; Underwood 1989).

3.1.2 Modes of presentation of recorded input

The nature of the input used in a PL task only accounts for one dimension of PL. To provide a full definition of a given PL task, one has to consider other features, such as the level of control that listeners have over the playback procedure, the actual playback sequence adopted, the nature of any additional input given to listeners, and the type of response that is expected from them.

3.1.2.1 Who has control over playback procedure?

Playback equipment may be controlled either by a third party (teacher, examiner, researcher) or by listeners themselves. Even in tasks using identical input, any change in the learners' degree of control over the flow of information and the choice of procedures can bring about significant differences in the learning process (Rost 1990:164; Wright 1987).

Controlled playback (i.e. playback that is controlled by a third party such as a teacher, examiner or researcher) is preferred in certain cases, on the grounds that listeners do not normally have the opportunity to play back the speech that they hear in real life in

[10] Unfortunately this is not always the case. For learners of certain FLs , the problem may be even more acute , due to culture-specific approaches to script writing in the media (e.g. Spanish TV news and documentary scripts tend to use a higher proportion of *literacy* features than the BBC).

the same way as they would play back a recording (Buck 1995; Ur 1984). Controlled playback is also the most commonly used procedure for testing purposes, although some listening tests give control to the candidates within a certain time limit (see section 1.2.3). Whenever controlled playback is used, it is common practice to tell listeners how many times a recording will be played before the task begins. Consequently they have some degree of choice in planning a listening strategy according to the playback sequence that is to be used. To that extent, they can be said to be using PL strategies. However, they are not able to apply playback strategies because they cannot alter the sequence of presentation themselves.

In contrast, *free playback* (playback controlled by the listener) does give her such choice. The fact that free PL tasks do not normally mirror listening tasks as they would occur in real life is not necessarily cause for concern. If one accepts Breen's proposed definition of "authenticity in terms of "authenticity of tasks conductive to language learning" (Breen 1985) "[11], PL can be regarded as an empowering device for the L2 learner. This is in line with the current interest in learner-based approaches aiming to promote learner autonomy (Campbell and Kryszewska 1992; Dickinson 1987; Dickson 1996; Peterson 1985; Wardrop and Anderson 1992). PL can also be regarded as a means to make the input comprehensible (Dupuy 1999; Krashen 1981; 1985; 1996), thus facilitating acquisition (see p. 75 for discussion).

3.1.2.2 What is the playback sequence used?

There is a potentially endless range of sequences that could be used in a PL task. These can all be defined in terms of three basic features:

1. The total number of exposures to the input (this may vary across different segments of the same recording).
2. The presence or absence of interruptions during each complete exposure
3. The type of output that is allowed/required (e.g. note-taking, ticking boxes, answering questions) during or after exposure.

[11] See previous discussion on p. 56.

The first feature will be discussed under "Repetition" (section 3.2.2.2 below); the second feature is related to pauses, which were discussed under "Temporal variables" (section 2.2.2 above). The last feature is discussed in section 3.1.2.4 below.

One of the aims of the present study is indeed to describe different combinations of repetition and pauses that listeners actually use in free playback, in order to categorise typical playback patterns (Experiment 2, section 8.3.4 below).

The effects of different sequences of presentation were tested by Sherman (1997) with a group of 78 intermediate-level undergraduates learning English in Rome. All subjects were tested under four conditions (all of them involving two uninterrupted presentations of the input): (A) with questions presented before listening, (B) with questions presented after the second playback, (C) with questions sandwiched between two hearings, (D) written recall without questions. After each test, students were given a questionnaire to elicit their views on the relevant procedure in terms of its methodological value, the amount of distraction, tension and effort that it generated, and the amount of extra information, support and freedom that it provided.

The highest test scores were obtained in the "sandwich" condition, and the lowest in the "no questions" condition. Conditions A and B (questions presented before / after input exposure) were not significantly different from each other. However, questionnaire results showed that students rated question preview as a "better"

	Advantages	**Disadvantages**
Condition A (Questions before)	Qs provide advance information (both propositional and schematic)	Qs prevent Ss from constructing their own interpretations of the global meaning (trying to make sense of the Qs while listening).
Condition B (Qs after two hearings)	Ss are free to construct their own interpretations of the global meaning while they listen.	Uncertainty [as to the nature of the questions] generates anxiety.
Condition C (Qs betwen two hearings)	1st hearing: Ss are free to construct their own interpretations of the global meaning. 2nd hearing: Qs make sense + provide advance information.	
Condition D (No questions: free recall)	Ss are free to construct their own interpretations of the global meaning.	["No Qs" means there is no way of monitoring one's performance]. Complete uncertainty generates the highest levels of anxiety.

Table 3-1: Qualitative effects of different playback procedures (adapted from Sherman 1997)

approach than presentation after exposure. "Questions after" was also perceived as the most effortful condition, as well as generating almost as much tension as the "no questions" condition, which was perceived as the most anxiety-generating of all.

Although subjects performed equally well under conditions A and B, they perceived B more negatively. The author concludes that question preview has a "high psychological value" for test takers, whereas "conditions of uncertainty, however cognitively favourable" are perceived negatively. Her discussion is summarised in Table 3-1 (notes in square brackets are my own).

Sherman's findings clearly show the relevance of playback procedure, both in terms of task/test performance and in terms of the listeners' perception of task/test difficulty. The main question addressed in this experiment was whether or not questions act as advance organisers when item preview is allowed. The discussion that follows will examine the role of advance organisers.

Rivenc proposes a series of teaching strategies that use playback in order to grade the amount of scaffolding given to the listeners (Rivenc 1995). A sequence comprising two approaches is suggested: the "slow" approach, which involves teacher-controlled playback of short segments accompanied by focused explanations, repetitions and questions, and the "fast" approach, in which longer passages are presented in real time for extensive listening tasks.

Studies examining PL from a learner-strategic point of view are reviewed in section 4.2.2 ("Strategy use in playback listening").

3.1.2.3 What cues is the listener given for schema activation?

In order to comprehend a message, listeners need to activate relevant schemata on which to base meaning hypotheses and inference-making (see p. 43). Right from the beginning of a task, learners will be actively (and often unconsciously) responding to any cues available for that purpose. In many learning situations, exposure to recorded input is preceded by some kind of *advance organiser*, which may take a variety of forms: a simple title, a short introduction (written or spoken), a picture, a vocabulary list, a task brief to be read/heard before listening, or a set of questions. Pre-listening questions may be based either on the subjects' previous knowledge, or on information to be found in the recording. The latter case is known in the testing literature as *item preview*. During exposure itself, new advance organisers may become available, for instance in the form of discourse markers or visual clues (if the

input is presented on video). Relevant studies investigating the effect of advance organisers on listening comprehension will now be reviewed.

The effect of advance organisers on comprehension and recall was tested in a frequently cited study by Bransford and Johnson (1972). In a series of experiments conducted in L1, they tested three types of contextual cues: a title indicating the topic of the message ("topic"), a picture showing elements present in the message, arranged in a way that did not reflect the message itself ("partial context"), and a picture showing the same elements in an arrangement that directly represented the message conveyed ("[full] context"). The experiments conducted were centered on three basic conditions:

1. No context (subjects just heard the recording, either once or twice).
2. Context given before listening (full or partial visual cue, or statement of the topic).
3. Context given after listening (full visual cue or topic statement).

The findings were very similar across the four experiments: both recall and comprehension were found to be significantly better in the "context before" condition than in each of the other conditions. In the case of visual cues, only the "full context" picture shown before listening produced significantly better performance.

In a study involving 200 Japanese college students, Buck studied the effect of item preview on their comprehension of English listening input (Buck 1990). A set of 33 items was presented to the subjects in two experimental conditions: with and without item preview. Buck found that item preview did not significantly improve performance, but did affect the items' discriminative power, producing a wider spread of difficulty among items in the "no preview" condition. In other words, the relative differences between "difficult" and "easy" items appear to be reduced when item preview is allowed. As part of a previous experiment in the same study, Buck analysed introspective reports from three subjects performing the same test with item preview. All three reported that item preview had helped their test performance (Buck 1990:328). Sherman's findings (discussed above) suggest that the potential benefits of item preview might be outweighed by the fact that it also focuses the listeners' attention on specific items in the input, thus preventing them from developing their own meaning interpretations. Nevertheless, item preview lowers anxiety and is therefore perceived as facilitating by the subjects, regardless of its actual effectiveness in cognitive terms (Sherman 1997). Both findings are consistent with Buck's results.

Berne (1995) compared the effects of different pre-listening activities and repeated exposure on 62 undergraduate learners of Spanish L2, randomly assigned to one of three conditions: question preview, vocabulary preview, and control group. Question items consisted of 10 multiple-choice questions in L1; vocabulary preview consisted of 10 key words, each with its L1 transation and an L2 example in context; the filler activity for the control group was to write down numbers from 1 to 50 in Spanish. The recording (a scripted lecturette read out on video by a native speaker) was played twice and responses to the MC questions were collected after each time. The questions preview condition obtained significantly higher scores than vocabulary preview and control group both times. The first time round, vocabulary preview scored even lower than the control group. This result (which failed to reach significance) was interpreted as an effect of attention being directed to form rather than meaning the first time round in the vocabulary preview group. Attention was assumed to switch onto meaning for the second exposure.

Teichert (1996) studied the effect of regular use of advance organisers and recorded input on L2 listening comprehension skills in German. The experimental group was exposed to video or audio input every week, preceded by a number of advance organisers, whereas the control group was not (the actual nature and quantity of listening input used in the control group is unclear). The author found that improvement on listening performance after a 14-week period of instruction was significantly greater for the experimental group. This, however, could be simply due to a greater amount of exposure made available to the experimental group (these subjects could listen to the recordings in their own time, unlike the control group). Therefore the results of this study cannot be regarded as conclusive.

For teaching purposes, the importance of advance organisers is generally assumed in common practice. Most listening manuals recommended that some kind of pre-listening activity is performed as a matter of routine prior to the first exposure to the input (Dunkel 1986; Grabe 1991; Mendelsohn 1995; Rixon 1986; Turner 1995; Underwood 1989).

3.1.2.4 What output is required from the listener?

A wide variety of possible response types can be found in listening tasks. Thompson (1995) lists the most typical formats that are used for testing purposes: they comprise selected responses (such as multiple choice and true/false questions), constructed responses (such as open-ended questions and recall protocols), other types of written

responses (such as supplying missing information in pictures, charts, diagrams, tables, forms etc.), and nonverbal responses (such as selecting pictures, graphs, drawing lines, arranging pictures). In some cases, listeners are required to respond physically by moving objects or performing simple actions such as getting up (Brown and Yule 1983b; Defilippis 1980). Other possible responses include oral repetition, dictation and even listening without giving a response at all for instance while following a written transcript or a sequence of pictures (Ur 1984).

In each case, the type of output that is required has a direct influence on the type of metacognitive and cognitive strategies that listeners use. A systematic discussion of all possible response types and their implications on listening strategy is, however, beyond the scope of this thesis.

In the previous discussion of Sherman's findings (Sherman 1997), it was reported that listeners who were required to produce a free recall felt more anxious and achieved lower test scores than those who answered a set of questions. Questions and free recall were also thought to elicit different types of processing (questions guide the listeners' interpretation, but they also restrict their freedom to develop their own meaning hypotheses). In general, listeners perform better and are more relaxed while listening if they know in advance what kind of response is expected from them, whereas uncertainty has the opposite effect (Anderson and Lynch 1988; Lund 1991b, cited in Thompson 1995:46; Sherman 1997; Ur 1984). Tasks requiring an immediate response are easier than those involving selection (such as summarising), which in turn are easier than those in which listeners have to distinguish between fact and opinion (Anderson and Lynch 1988).

It must be noted that, within each of the categories discussed above, different levels of processing may be possible. Therefore, the strategies used depend not only on formal aspects of the response (e.g. multiple choice vs. open-ended questions), but also on the nature of the information required as a response. Shohamy and Inbar (1988) examined the effect of question type on tests of listening comprehension, using three question types: (1) global understanding; (2) local understanding; (3) trivial information. Performance was significantly better on "local" questions than on "global" questions, and this occurred across all topics, tests and proficiency levels used in the experiment. The results obtained for "trivial" questions were inconclusive.

In order to reach a full understanding of the strategies involved in PL, future research will need to examine how these operate in a variety of task types. This study focuses

on two of the most commonly used responses (open-ended questions and free recall). Any playback strategies identified for these two task types, however, cannot be assumed to be extendable to PL tasks requiring different response types.

3.2 PL as a specific skill

Listening to recorded input with control over playback entails certain skills that are specific to this modality, and differ in some ways from the skills involved in listening tasks in other modalities (often known as "real-life" listening). For convenience, these two categories will be referred to as *playback listening* (PL) and *naturalistic listening* (NL) respectively. This section discusses which specific features of PL may hinder or facilitate comprehension and learning in comparison to NL. The discussion below focuses primarily on PL that uses audio (ARI), rather than video-recorded input.

3.2.1 Adverse features for L2 learners

3.2.1.1 Lack of visual channel

The first disadvantage of audio-based PL is that audio recordings deprive the receiver of the visual channel (Kellerman 1990).

In NL situations in which visual communication is not available, speakers make certain adjustments to compensate for this. Boyle et al. (1994) observed the interactions of task partners who were deprived of visual contact. Lack of non-verbal information caused these subjects to use more back channel responses[12] in order to increase verbal feedback to their task partners.

Vision also provides redundant information on the articulatory movements made by the speaker. Kellerman (1990) reports experiments in which subjects were exposed to conflicting messages from sound and vision channels, such as lips shown pronouncing [ba], but audio-dubbed as [ga]. This produced consistent misperceptions of [ga] as [da]. Such shift of perception to an intermediate place of articulation indicates that listeners do process both visual and auditory information at the sensory perception stage.

Kinesic behaviour (body movements) is another aspect of visual information that can aid comprehension at several levels: gestures provide phonological and semantic

[12] Feedback that listeners give to speakers in order to indicate success of failure in communication (e.g. headshakes, "uh", "really", "yeah", etc.).

reduncancy, operate alongside discourse markers to signal the structure of the spoken message (Levy and McNeill 1992), and give signals to regulate interaction - e.g. turn-taking, listener feedback (Kellermann 1992). Visual clues also act as advance organisers: they help viewers activate relevant schemata in order to interpret the situational context of the message. (Allan 1985; Willis 1983).

Given the wealth of information that vision contributes, it can be concluded that, compared to ARI, video elicits listening skills that are closer to those used in NL (Rubin 1995). Indeed, Kellerman (1990) found performance on video-based listening comprehension to be a better predictor of NL comprehension than audio-based performance. Parry (1984) compared the performance of 178 English speakers in listening comprehension tasks using audio and video-recorded input in Spanish. The audio version was a copy of the soundtrack from the video (i.e. auditory input was identical in both conditions). The video condition produced significantly better performance. The use of video over non-visual media is widely supported in the literature, not only on the grounds of improved comprehension (Herron et al. 1995; Herron and Seay 1991; Rubin 1990), but also improved retention (Rubin 1995), especially in a multimedia context (Brett 1997). However, it has been pointed out that such benefits on retention have not as yet been confirmed in the long-term (Lynch 1998:7). It is also worth noting that the relationships between sound and vision are not necessarily always redundant, or even complementary (Lancien 1986), and that in some cases (e.g. newsreaders on autocue) the visual channel may even be distracting (Rubin 1995).

3.2.1.2 Quality of the acoustic signal

Another source of difficulty is that noise and signal are mixed together and therefore even more difficult to discriminate in ARI than in NL. Using stereo recordings may help in this respect, because stereo contains spatial information, that is not present in mono. A personal anecdote may be relevant here. A few years ago, one of the mature students on the course I was teaching had mild hearing difficulties and was particularly uncomfortable with language laboratory work. Eventually, she brought a letter from her medical consultant, recommending that we used stereo recordings wherever possible. Stillman (1980, quoted in Garman 1990:21), explains how the "stereo" effect creates a sense of space by comparing the signals received from right and left ear. Just like the signals from our two eyes are combined to create the

stereoscopic image that enables us to appreciate distances visually, signals from our two ears are combined to create the stereophonic sound that enables us to appreciate from which direction each of the sounds we hear is coming. If this feature is removed, the brain will have no sensory clues to discriminate the voice of one speaker from background noise or from other speakers (Drullman and Bromkhorst 2000), and this will impose a greater load on cognitive processing. If video is used instead of audio, the difficulty may be partly compensated (provided that the relevant speaker is in shot), as the visual signal will be filling in some of the missing information (e.g. lip movement).

3.2.1.3 Lack of listener-speaker interaction

The most criticised disadvantage of PL is that there can be no interaction between speaker and listener. In collaborative discourse, listeners play an important role in shaping the interaction in a variety of ways (Long 1983; Rost 1990; Rost and Ross 1991). Listener responses in collaborative discourse can be summarised as follows:

Skilled listeners will...

1. Try to identify points at which they can switch to speaker role.
2. Organize turn-taking and provide obligatory responses.
3. Provide "listenership cues" to show the speaker that they are following.
4. Prompt the speaker and indicate which aspects of the discourse are to be developed.
5. Provide cues to indicate how they align with the speaker's claims.
6. Reformulate the speaker's contribution when it conflicts with their listener goals
7. Identify a plausible intent on the speaker's part when interpreting an utterance.
8. Coordinate speaker and listener purposes in the discourse.
9. Query and repair when appropriate.
10. Check their understanding when appropriate.

(adapted from Rost 1990:116)

When interacting with non-native interlocutors, native speakers spontaneously adjust their discourse to the listener's proficiency (Lynch 1989). In a study of interaction between native and non-native speakers, Long (1983) also found that modifications of interaction between a speaker and listener (by conversation checks, expansions, repetitions and clarification) facilitate comprehension more than language modifications made by the speaker independently of the listener.

None of the above is possible when listening to ARI.

3.2.2 Facilitating features for L2 learners

3.2.2.1 Listening anxiety

Given that in listening, much of the processing load is due to the constraints of on-line processing, an opportunity to hear the input again ought to reduce a listener's anxiety, especially if the subject has direct control over the playback machine.

Language anxiety research is primarily concerned with speaking anxiety, possibly because there was already more research on this area in L1 (Horwitz and Young 1991). Talking in public is perceived as more problematic than listening in one's native language, given that *conspicuousness* is a major component of communication apprehension (Bacon and Finnemann 1990; Daly 1991). There is surprisingly little reference to the anxiety that time-related pressure may produce in L2 listening tasks (Vogely 1998). Horwitz and her colleagues (Horwitz, Horwitz, and Cope 1991) developed a Foreign Language Classroom Anxiety Scale (FLCAS), comprising 33 questions. Out of those, only 3 questions relate to listening, of which 2 refer specifically to listening in the foreign language (both refer to listening to the teacher, none to recorded input). In contrast, there are 8 questions related to speaking anxiety, 5 of which directly refer to speaking in the L2. A few other scales have been developed to assess listening anxiety in greater detail, but all of them were designed with L1 listeners in mind (Beatty, Behnke, and Henderson 1980; Spielberger, Gorush, and Lushene 1968; Wheeless 1975).

The most informative study on L2 listening anxiety so far was conducted by Vogely (1998). It took the form of a questionnaire that was given immediately after the listening section of an examination paper (recordings played twice, answering comprehension questions in the traditional format). Subjects were 140 students enrolled on undergraduate Spanish courses. The questionnaire consisted of three simple questions: "Do you experience anxiety when you are participating in a listening comprehension activity?"; "What makes you anxious when you are participating in a listening comprehension activity?" ; "What types of exercises, settings, or activities help to lower your anxiety level?".

Anxiety related to the input accounted for 51% of the problems reported: nature of the speech (e.g. too fast); difficulty (of vocabulary, syntax, topic); lack of clarity as to what response is required (this is consistent with Sherman's (1997) findings); lack of visual

support; lack of repetition ("two strikes and you're out"). Suggested solutions included: practice with audio material and repeated playback ("repeat the passage, repeat the passage, repeat the passage", student's reported wording).

Process-related problems accounted for 30% of the responses: inappropriate strategies (Ss perceive listening comprehension as a matter of word-for-word decoding); lack of processing time; not knowing how best to prepare for listening comprehension; inability to check answers. In contrast with the large proportion of problems reported in this area, only 4% of the suggestions made were in his category. This suggests a clear need for strategy training (acknowledged by Ss themselves). Other suggestions included being allowed to respond to the questions in L1. No explicit mention of student-controlled playback occurred as a possible solution.

Instructional factors accounted for 6%: lack of listening practice; "the test thing"; physically uncomfortable environment. A disproportionately high number of solutions (60%) were suggested for this type of problem, including more practice in class, regular feedback, combining listening with other skills (listen and repeat, read a transcript); opportunities for genuine interaction with native speakers.

Finally, personal attributes of teacher and learner accounted for 5%: fear of failure/nerves related to past experience; intimidating instructor. Solutions proposed included being given opportunities to experience small successes, and the use of anxiety-reducing techniques such as breathing exercises.

Among the recommendations made by the author is the use of clear and concise structured tasks (in cyclic rather than linear progression), using texts "recurrently" and focusing on "layers of comprehension", gradually moving from controlled output to more open-ended responses. She also stresses that teachers should understand the constraints of the listening process, in which the listener has no control over the rate of processing (Vogely 1998: 75). Vogely's findings confirm that a substantial proportion of anxiety-generating factors in a conventional listening test is indeed related to lack of control over the playback procedure. Reading the sample of *verbatim* responses reported in the paper, it becomes clear that students are explicitly making a plea for greater use of repeated playback and would emotionally (and perhaps also cognitively) benefit from greater control over the entire playback procedure, both in testing and practice situations.

3.2.2.2 Repetition in PL

Learners who are given control of the tape can get back to difficult parts of the message, just like readers do. In an interactive situation, the nearest real life equivalent to tape playback is repetition that is prompted (explicitly or implicitly) by a request from the listener. The difference is that a tape will play back identical input each time, whereas a speaker will tend to adjust the message to the perceived needs of the listener (Long 1983), or at least produce a slightly different message from the original one (Field 1997). Identical repetition is a distinctive feature of tape playback. On the one hand, repetition of this kind is of no use when the listener is simply unable to understand the utterance in the first place. On the other hand, repeated exposure to identical input has beneficial effects of its own.

Repetition increases redundancy, which should facilitate comprehension. Long (1991) analysed introspective reports from 6 college students of intermediate Spanish after listening to video-recorded input. One of the strategies reported was "focusing on redundancy". Other introspective studies of listening strategies mention repetition as a self-reported strategy (Laviosa 1990; 1991b). Whenever subjects are allowed to use repetition, they generally make use of it for strategic purposes (Defilippis 1980; Merlet and Gaonac'h 1995; Thompson and Rubin 1996; Zhao 1997).

Other authors have stressed the pedagogical importance of repeated exposure to identical input using different set tasks each time (Lund 1991a; Weissenrieder 1987). Lund (*ibid.*) also reports that repetition appears to be more beneficial to listeners from intermediate level upward.

The benefits of repeated exposure have also been found to be more significant in listening than in reading. Hildyard and Olson (1982) found that the differences in recall that could be observed between reading and listening (less detail information in the listening recalls) decreased with a second exposure to the input: more detail information was gained by both readers and listeners after the second exposure.

Tyler (1994) found that certain patterns of repetition in the spoken input (e.g. lexical repetition) contribute to comprehensibility, and that absence of such patterns results in perceived incoherence. However, as Thompson (1995:39) points out, all forms of redundancy are not equally helpful to L2 listeners:

Chaudron (1983) compared the effects of 4 types of topic reinstatement devices on topic identification and recall. The devices tested were: simple noun ("The beer tastes terrific"); synonym ("The brew tastes terrific"); repeated noun ("The beer... the beer

tastes terrific"); topicalising rhetorical question ("What about the beer? It tastes terrific"); non-conditional if-clause ("If you can afford the beer, it tastes terrific"). Redundancy in the form of a repeated noun proved the most helpful device of all syntactic modifications tried out, particularly for low proficiency subjects. This finding suggests that the type of redundancy that PL provides should potentially have beneficial effects on comprehension, because identical repetition allows listeners to hear the same words again (the most helpful type of redundancy). In contrast, redundancy that uses not only repetition of constituents, but also paraphrase and synonyms has been found to help only L2 listeners with higher levels of proficiency (Chiang and Dunkel 1992)

In a study that was discussed in section 3.1.2.3 above, Berne (1995) found that repeated exposure to the recording was more effective than advance organisers such as question preview or vocabulary preview in improving listening comprehension performance. All groups (question preview, vocabulary preview, and control group) performed significantly better the second time round. Her findings also suggested that repetition was most beneficial to subjects in the vocabulary preview condition, who doubled their average score the second time round (although the interaction between pre-listening activity and exposure failed to reach significance level). Berne concludes that repeated exposure is so helpful that it ought to be considered in the design of listening comprehension tests (Berne 1995).

Cervantes and Gainer (1992) studied the listening performance of 82 Japanese learners of English. Performance was measured with a partial dictation test, in which 3 to 7 words had been deleted from 13 separate dictation segments on the transcript of a lecture. Three conditions were compared: (1) "simple" version of the lecture played once, (2) "complex" version of the same lecture played once, (3) the same "complex" version with each dictation segment played twice. Performance was significantly better in condition 1 than in condition 2, but there was no significant difference between versions 1 and 3. This suggests that repetition may compensate for syntactic complexity, at least within the very specific parameters used in this experiment (repetition by segments in a partial dictation exercise). Due to a sampling error, the authors were unable to assess level of proficiency as a variable.

Identical repetition of recorded input allows listeners to apply selective attention strategies and focus intentionally on specific segments. Within a particular segment, each new listening allows decoding of additional units, as the sets of units decoded in previous runs do not need to be processed at low level anymore. This enables

listeners to allocate more resources to the processes that could not be completed the first time round. According to capacity theories, such increase in the resources available ought to result in better performance (Just and Carpenter 1992; Zwaan and Brown 1996). When a message is repeated in identical form, no new information is introduced that might create new processing needs. This could explain why verbatim repetition is effective at low proficiency levels (Chaudron 1983) while other forms of redundancy are not (Chiang and Dunkel 1992).

Repeated playback of recorded input is also seen as beneficial in terms of language acquisition. Krashen proposes the concept of "narrow" input, which consists of selecting from all input available those sources of "mainstream" (real-life) input that are most likely to provide input at the critical *i+1* level (Krashen 1985). Later developments of this principle propose its application to specific skills under the names of *narrow reading* (Krashen 1981) and *narrow listening* (Krashen 1996). Narrow listening is a learning technique in which the learner records several native speakers talking for 2-3 minutes about a topic of the learner's choice. Listening to those recordings again and again is claimed to result in improved comprehension and L2 acquisition. A survey conducted by Dupuy (1999) concluded that intermediate college FLE students exposed to narrow listening found it particularly useful for improving comprehension, fluency, vocabulary and confidence with the L2. Subjects were asked to report the number of times that they had played each recording and what percentage they had understood after the first and last time. Considerable gains (20% or more) were reported after repeated exposure. As for the number of repetitions, the majority (70%) of semester 1 and 2 students listened twice to each speaker on average, whereas the majority (66%) of semester 3 and 4 students reported an average of 3 times. It must be pointed out that prior to the study, all subjects had been given a handout explaining the rationale of narrow listening, in which the following statement was made:

> Previous investigations show that, for beginners, seven listenings of the same segment is a good average to understand what is being said. With every additional listening you do, your understanding will increase by 10% on the average (Dupuy 1999:360).

This, in addition to the fact that the results were based solely on self-reported percentages (a highly subjective measure) suggests that Dupuy's claims should be tested against more reliable evidence before any assumptions on the value of narrow listening can be made.

In order to compare the effects of repetition and speech rate on listening comprehension, Zhao (1997) used a computer programme that allowed listeners to control speech rate electronically. He found that, when subjects are able to set speech rate to their own "ideal" speed, the effects of speech rate adjustment on comprehension are more beneficial than those of repetition (see p.84 for a more detailed account of Zhao's experiment). Subjects who were able to use both devices relied more on speech rate adjustments than repetition. Nevertheless, 21 subjects out of the 30 whose experimental condition allowed repetition did make use of it at some point.

4. Listening strategies

This chapter discusses the concept of "strategy", first with reference to current research on general strategies for listening skills, then more specifically in relation to playback listening. Strategy training is discussed in section 4.3. The final section discusses commonly used methods in strategy research.

4.1 Defining "strategy"

Tarone (1981, quoted in O'Malley and Chamot 1990:43), distinguishes four types of strategy: learning, communication, production and perception strategies.

1. *Learning strategies* reflect the desire to learn the target language (actual communication is not the main focus).
2. *Communication strategies* are used for negotiating meaning when communication fails.
3. *Production strategies* are attempts to use the language system efficiently in order to communicate a message.
4. *Perception strategies* are attempts to interpret utterances efficiently.

The distinction between these categories is not always clear-cut because these strategy definitions are based on the subject's intention, which is rarely explicit (O'Malley and Chamot 1990; Tarone 1981).

- *Production strategies:*

Production strategies may occasionally be used in PL situations in which listeners are performing a task that integrates several skills (e.g. problem-solving simulation). However they will not be discussed here, as productive skills are not the focus of the present study.

- *Communication strategies:*

Kasper and Kellerman (1997) consider two possible aspects of communication strategies: these can be regarded as "intraindividual" (the psycholinguistic view) or "inter-individual" (the interactional view). Inter-individual communication strategies have already been mentioned in section 3.2.1.3 (p.70), where it was noted that this

type of strategy cannot be used in PL (For further studies of communication strategies, see Bialystok 1990; Kasper and Kellerman 1997).

- *Perception strategies*

Naturalistic listening normally elicits perception strategies: in typical real life situations, the main purpose is to comprehend the message for transactional or interactional purposes. In contrast, PL by an L2 listener is more likely to be related to learning strategies: the ultimate purpose is to gain communicative competence. Of course, a few real-life situations involve PL for purposes other than language learning (see section 1.2.1, p.14); and reciprocally, an L2 listener may well be engaging in language learning while performing a transactional task (Tarone 1981). Once again, the fact that the definition is based on the listener's intentions makes it difficult to determine to what category a particular strategy may belong.

Tarone (1981) defines perception strategies as "attempt[s] to interpret incoming utterances efficiently, with the least effort", of which she gives examples such as "pay attention to the ends of words" or "pay attention to stressed syllables".

The current definition of perception strategies may be incomplete and problematic in terms of its reliance on subjective notions such as the receiver's "intention". However, the fact that perception strategies are defined as a distinct category from learning strategies is in itself conceptually relevant, regardless of any practical difficulties in actually discriminating one group from the other.

- *Learning strategies*

According to Weinstein and Mayer (1986, quoted in Chamot 1995:13), learning strategies are "the steps, plans, insights, and reflections that learners employ to learn more effectively. Learning strategies are intentional on the part of learners, and are used for the purpose of facilitating learning."

Oxford (Oxford 1990:8) defines learning strategies as "specific actions taken by the learner to make learning easier, faster, more enjoyable, more self-directed, more effective, and more transferrable to new situations". According to her, the features of language learning strategies are that they...

1. Contribute to the main goal, communicative competence
2. Allow learners to become more self-directed
3. Expand the role of teachers
4. Are problem-oriented
5. Are specific actions taken by the learner
6. Involve many aspects of the learner, not just the cognitive
7. Support learning both directly and indirectly

8. Are not always observable
9. Are often conscious
10. Can be taught
11. Are flexible
12. Are influenced by a variety of factors

One difference to be noted between the two definitions is that Weinstein and Mayer's definition (1986) states that strategies are "intentional" steps taken for a "purpose", whereas Oxford merely claims them to be "specific actions" that are "*often* conscious" and "*contribute* to the main goal" [my emphasis]. Even if we assume the terms "conscious" and "intentional" to be synonymous, there remains an issue in the degree of consciousness/intentionality that strategies imply.

Rabinowitz and Chi (1987:84) explicitly narrow down their definition of "strategies" to those processes that are "goal-oriented" and "intentionally invoked". This implies that a mental process that takes place without conscious awareness cannot be regarded as strategic behaviour. The position supported by O'Malley and Chamot is somewhat more open: the authors accept the distinction between processes and strategies insofar as "strategies that occur overtly cannot qualify as mental processes", but they also argue that "nevertheless, the mental processes underlying these overt strategies could easily entail such strategic modes of processing as self-monitoring, summarizing, and inferencing (among other strategies)" (O'Malley and Chamot 1990:88). This consideration could also be applied to describe the complex relationships existing between automatic processes and controlled strategies in PL.

The common assumption of consiousness in strategy definitions has also been questioned by Bialystok (1990), who pointed out that strategies are often defined as instances of conscious behaviour, but this may not necessarily be so: "Communication always involves choice, and the choices evident when a strategy has been used may have been made no more or less consciously than any other choice" (Bialystok 1990:4). She also questions the assumption whereby all strategies are intentional.

The difficulties in reaching a convincing definition of the term "strategy" seem closely related to their reliance on the notion of consciousness (either in order to include it or exclude it from the notion of strategy), a matter that was pointed out by McDonough (1999) in a recent review of strategy research. The use of consciousness (or lack of it) as a criterion is problematic because the term is often based on subjective notions

Schmidt (1990; 1994) offers a possible way out of this problem, by deconstructing "consciousness" into four other notions that are commonly (but not consistently) related to it by researchers. He explains that in the language learning literature, the term "consciousness" can be found to mean:

1. *Intentionality* (intentional versus incidental learning)
2. *Attention* (focal attention and "noticing" versus peripheral attention)
3. *Awareness* (explicit versus implicit learning)
4. *Control* (controlled versus automatic processing).

Two possible scenarios for strategy use illustrate how complex the interactions between different notions of "consciousness" can be:

- A strategy may occur as a result of attention being involuntarily drawn onto a salient item in the input (Schneider and Shiffrin 1977; Shiffrin and Schneider 1977).
- A particular strategy may involve a mixture of controlled and automatic processes operating together (Rabinowitz and Chi 1987).

The difficulties of current research in reaching a widely accepted position regarding the role of consciousness in strategic behaviour suggest that there are potential dangers in operationalising the notions of *strategy* and *process* in terms of conscious vs. unconscious behaviour. The present study therefore chooses to avoid any specific focus on consciousness as a determining feature of behaviour. Instead, it will focus on any processes or strategies (conscious or otherwise) that result in overt behaviour. In this operational context, an explicit distinction between strategy and process becomes of little relevance.

4.2 Learner strategies for L2 listening

4.2.1 Identifying and classifying listening strategies

The most commonly used classification of learning strategies considers three different categories, depending on the type and level of processing involved: metacognitive, cognitive and social mediation/social-affective (Brown and Palincsar 1982; O'Malley and Chamot 1990; O'Malley et al. 1985b).

- *Metacognitive strategies* focus on the executive aspects of the task, and comprise selective attention, planning and monitoring, and evaluation.
- *Cognitive strategies* focus on incoming information itself, and comprise: rehearsal, organisation, inferencing, summarising, deducing, imagery, transfer, and elaboration.
- *Social-affective strategies* cover a range of strategies centered on interaction with other people or the management of one's own emotions while completing the task. They comprise: cooperation, questioning for clarification and self-talk.

(O'Malley and Chamot 1990:46):

Inspired from Rubin's (1975) list of things that "the good language learner" typically does, Oxford (1990) considers the categories listed above within a wider distinction. She proposes two broad categories: direct strategies (involving direct learning of the new language) and indirect strategies (contributing indirectly to learning). In her taxonomy, cognitive strategies are regarded as direct strategies, along with memory and compensation strategies; whereas metacognitive, social and affective strategies constitute the group of indirect strategies. Each group is subdivided into three further levels, comprising altogether a list of 62 different strategies, that covers strategies as different from each other as "highlighting" (cognitive strategy) and "listening to your body" (affective strategy).

In recent years, a number of studies have been conducted in order to identify the skill-specific strategies that L2 learners use in order to improve their listening skills and/or comprehend specific spoken messages.

O'Malley and his colleagues (O'Malley and Chamot 1990, studies 1 and 2; O'Malley et al. 1985b) conducted a series of exploratory interviews with groups of learners studying Spanish and Russian at beginner and intermediate level. The groups were prompted to describe the strategies that they used in a number of classroom and out-of-class activities such as vocabulary learning, oral and written grammar drills, listening comprehension, reading comprehension, written composition, oral presentations, social communication (e.g. conversation with a NS), and operational communication (e.g. ordering a meal). The strategies identified are shown in Table 4-1.

To narrow down these findings to specific listening strategies, the authors conducted a series of talk-aloud sessions (O'Malley, Chamot, and Kupper 1989; 1990, study 3) in which ESL students performed specific listening tasks, using recordings of a history

A. Metacognitive strategies:

Planning

- Organizational planning
- Delayed production

B. Cognitive strategies

- Rehearsal
- Translation
- Note-taking
- Substitution
- Contextualization

C. Social/Affective strategies

- Self-talk

Table 4-1: Learning strategy classification
(O'Malley and Chamot 1990, study two: p.126)

lecture, a science lecture, a short story and a dictation passage. The recordings were paused at regular intervals, during which the researcher prompted introspection with simple probe questions such as "What were you thinking?". The strategies identified in this way were classified according to Anderson's (1985) three phases of listening (perceptual processing, parsing and utilisation) as shown in Table 4-2.

Perceptual processing:	• Selective attention • Self-monitoring
Parsing	• Grouping (listening for larger chunks) • Inferencing from context
Utilisation:	• Elaboration from world knowledge, personal experiences, or self-questioning

Table 4-2: Listening strategy classification
(O'Malley, Chamot, and Kupper 1989; 1990, study 3)

In a longitudinal study of 19 learners of Russian or Spanish (O'Malley and Chamot 1990, study 4), O'Malley and Chamot recorded the subjects' reported strategy use

over four semesters for four task-types: vocabulary, listening, cloze, writing. The subjects' preferred strategies for listening tasks were:

- Metacognitive strategies:
  *selective attention, self-monitoring and *problem identification.
- Cognitive strategies:
  *note taking, elaboration, inferencing and summarising.

Strategies marked with an asterisk are those that were exclusively reported as "preferred strategies" for listening tasks. The study did not compare specific strategies used across different listening tasks.

The only longitudinal data available for listening was from 4 learners of Spanish (all of them identified as "effective" at the beginning of the study). These students were found to plan less at the end of the study than at the beginnning. They also reported more self-monitoring, elaboration and summarising strategies than at the early stages.

4.2.2 Strategy use in playback listening

There are few specific references to PL in the literature, but a number of studies on listening strategies mention strategies that are more or less implicitly related to PL. The most commonly mentioned PL strategies are those involving repeated playback.

One of the methods reported is to allow subjects to play the tape on their own, and then ask them how many times they played the recording (Chien 1998; Dupuy 1999). Retrospective reports may indeed be appropriate in certain cases for studying covert listening strategies. However it seems a unfortunate to rely on self-reported behaviour when the behaviour in question is one of the few strategies that are readily accessible to external observation (i.e. pressing buttons on a playback machine).

A more common method is the use of recorded input to elicit introspective talk-aloud. The tape is paused at regular intervals, during which the researcher prompts introspection (Buck 1990; O'Malley, Chamot, and Kupper 1989; O'Malley and Chamot 1990; Vandergrift 1998). Other researchers prefer to collect data only once the recording has been played straight through (immediate retrospection) in order to avoid interfering with the listener's natural strategies while performing the task (Bacon 1992a; 1992b; 1992c).

A few studies use computer technology, either in order to log the subjects' playback strategies, or for training purposes. Cauldwell (1996) describes computer software in which particular realisations of spoken words can be isolated by means of non-destructive editing and made available to L2 listeners for observation and learning. In one of her studies, Bacon (1992b:328) reports an original strategy that shows a listener's attempts to use hardware features in a similar manner: the subject in question used the volume monitor to assist segmentation "I just watched the lights go up and down and that gave me an idea of where the words separated". Such intuitive use of technological features for listening purposes is far removed from the strategies that listeners would use in real life. However it does show that learners will spontaneously devise whichever listening strategies make sense to them at the time in order to perform the task in hand. To achieve their goal, they are prepared to utilise any resources available, whether these are natural or engineered.

Zhao (1997) used a computer programme that allowed listeners to control speech rate electronically. Four experimental groups were given varying levels of playback control. The four conditions are summarised in the table below:

Subject's control over playback:	**C1**	**C2**	**C3**	**C4**
• Speech rate setting before listening [using a mock recording as warm-up]	No	Yes	Yes	No
• Speech rate adjustments while listening	No	No	Yes	No
• Repetition	No	No	Yes	Yes

Table 4-3: Levels of subject control over playback (Zhao 1997)

In terms of playback strategies, the following results were obtained:

- Condition 2: 14 out of 15 subjects used slower speech rate than preset.
- Condition 3: 8 out of 15 varied speech rate 2 or 3 times during the task.
  7 out of 15 remained at a slower speech rate than preset.
- Condition 3: 8 out of 15 subjects repeated input once.
- Condition 4: 13 out of 15 subjects repeated input between 1 and 3 times.

In other words: all but one subject (29 out of 30) made use of whichever control they had over speech rate. Of those whose only control was over the number of repetitions, all but two (13 out of 15) made use of such control. This indicates that,

given control over playback, the vast majority of L2 listeners use it strategically in order to improve comprehension. Given a choice, control over speech rate is used more often than control over repetition. Such preference was consistent with listening comprehension scores, which were higher in condition 2 than in condition 4.

Long (1991) used the talk-aloud technique with 6 learners of Spanish while they worked through a video recording with full control over playback. Their set task was to answer five comprehension questions. Note-taking was allowed and subjects "were encouraged to pause, rewind and listen as many times as necessary". Subjects made frequent references to the difficulty of processing all the information from both audio and video channels at once, even though the pictures were said to convey helpful information (see also Coniam 2001; Merlet and Gaonac'h 1995). Long identified the following strategies [my comments in square brackets]:

1. Paying attention ("don't stop listening if you get stuck")
2. First, play through the whole recording once. "All six subjects elected to play the video through completely before going back and relistening to the passages that they identified as segments (those containing answers to the detail question)". [Long interprets this as an organising or orienting strategy].
3. Recognising key words (the most frequently verbalised strategy).
4. Utilising background knowledge (with or without success).
5. Taking advantage of text length [the author's meaning is not very clear: waiting for supplementary cues later on in the recording?]
6. Utilising pauses [but Long seems to refer to natural pauses that are already present in the input rather than playback interruptions made by the listener].
7. Focusing on redundancy (e.g. interpreting enumerations)
8. Taking notes (very few did) [perhaps because this is more problematic with video than it would be with audio (cf. Coniam 2001)]

The author's conclusion regarding PL is that "allowing subjects complete control over manipulation of the text revealed that they use a multi-layered processing strategy, watching first for visual cues, then listening for audio cues" (Long 1991). However, the reported strategy sequence is surprisingly neat, and would have to be taken at face value: no specific evidence is given from protocol data directly related to the task being performed in the study. One of the subjects did describe a previous listening experience (unrelated to the study) in which she had used repeated

exposures in order to acheive comprehension. She reported watching the Argentinian film *La historia Oficial* three times. The first time she focused on reading the subtitles (and felt unable to watch the pictures), the second time she watched the pictures, and the third time she finally reports "understanding" the film by combining information from all three channels: subtitles, pictures and speech (Long 1991:13).

In a case-study involving 7 learners of French, Vandergrift also mentions "strategies related to repetition" (Vandergrift 1998:376+390). The researcher makes repeated mentions of a "second time", although the precise number of exposures to the input is not specified in the reported procedure. One of the cases described is a subject who...

> ...has a clear notion of the content schema which she has verified throughout her first listening effort. Furthermore, she has identified areas that need further clarification. She now has the requisite background information and planning strategies to direct her second listening effort and make it as productive as possible. (Vandergrift 1998)

In terms of PL strategy, this example could be interpreted as follows:

1. *First hearing*: activating a plausible schema (no context information was given prior to listening) and locating specific information (without attempting to decode it yet).
2. *Second hearing*: filling in slots in the schema by decoding the specific information that has been located during the first hearing.

Defilippis (1980) found that unskilled listeners requested overall more hearings than skilled listeners. He identified 19 listening strategies, two of which are directly related to PL:

> Strategy no. 10 - *Particularizing*: "The listener, during successive repetitions of the auditory material, listens expressly for particulars or details which were not understood during the initial playback"
>
> Strategy no.14 - *Repeated listening*: "The listener explicitly states that repeated playbacks of the auditory stimulus were used as an aid in processing the aural comprehension tasks"
>
> (Defilippis 1980:85-86)

Strategy no.10 would appear to be a subset of strategy no.14 (a more rigorous approach to strategy definition would seek to avoid this kind of overlap, see section 4.4.3). It must be pointed out that repetition was also mentioned occasionally as one of the "frustrating" aspects of the task (same speech rate, same wording every time).

These results of the Defilippis study are consistent with the most commonly claimed arguments in favour and against repeated listening.

Laviosa (1993) mentions two PL strategies reported by advanced learners of Italian: focusing on speakers' voices and background noise (for clues from intonation and context-related noises); and another strategy to which she refers as "listening backward":

> After repeated listenings, students can reconstruct the meaning of the message by combining information heard at the end of the segment with other pieces of information contained in the message (...). They reach the global meaning of the message through a combination of intermediary steps that move back and forth and that also proceed from the end towards the beginning of the message. (Laviosa 1993:300).

In a longitudinal strategy-training programme conducted with learners of Russian in a US university, Rubin and Thompson (1992) identified a number of strategies that subjects used when listening to video-recorded input (Table 4-4.)

Listen to: cognates international words key words and phrases proper names familar words partially familiar words repeated words and phrases tone of voice intonation	Notice the logical structure of the clip: logic of the storyline logic of the action logic of relationships Goal definition: look for basic schema attend to specific words/phrases decide when enough is enough
Prediction (words, sentences or schema), using: visual information general word knowledge information from clip itself	Action plan: listen with sound on or off determine how many times to watch break up into portions
Determine genre of the clip: news report interview dramatic episode commercial	Monitor verify prediction(s) define source of difficulty jot down problematic portion(s) consider other strategies
	Resources: discuss with classmate ask teacher

Table 4-4: Strategies used by learners of Russian in the Rubin and Thompson's longitudinal study (Rubin and Thompson 1992)

From the findings of this study, the authors concluded that playback strategies such as deciding when to use sound on/sound off or deciding when to play back again can indeed be taught. Studies on strategy training will be discussed in section 4.3 below. Thompson and Rubin's list also shows that the nature of the input presented in the recordings is thought to have a direct influence on the strategies that listeners use.

Bacon (1992a) studied how listeners alter their comprehension strategies between less difficult and more difficult passages. Her subjects were 50 university students (3 classes), studying Spanish at intermediate level. They were presented with two fully scripted authentic recordings at two levels of difficulty which were determined mostly in terms of topic familiarity (although speech rate was also slightly faster in the "difficult" recording). Retrospective interviews conducted immediately after the listening tasks showed no statistical difference between the proportion of top-down to bottom-up strategies used in the two recordings. Overall, students appeared to have little awareness of metacognitive strategies. Furthermore, whenever a greater number of while-listening metacognitive strategies was reported in the "difficult" text, comprehension seemed to suffer. Only 14%-18% reported using advance organisers before the task, and less than 1/3 evaluated their comprehension or strategies after listening. "For many, the task seemed to be over once the tape had ended" (Bacon 1992a:409).

As well as the input presented, task format may also have an effect on strategy use, as seen in Sherman's study on item preview (Sherman 1997, see p.63).

Defilippis (1980) studied the strategies used by 38 college students of French when performing five different listening tasks (see Table 4-5 below). He concluded that "the instructions given to listeners play an extremely important role in so far as they alert listeners to certain demands of the aural comprehension tasks", thus causing them to employ strategies that they perceive as "consonant with task demands" (Defilippis 1980:117). The specific strategies that were identified in relation to each task type have been summarised in Table 4-5 below (numbered lists in the right-hand column indicate strategies that are applied in temporal sequence).

Task description	Strategies reported
1. **Identification:** [2 to 4 hearings, as preferred] (1.a) match statements to one of 4 pictures, shown a few secs. after hearing each set ("Après le dîner, ils prennent du café") (1.b) state whether a sentence was a statement or a question ("Vous Dînez chez les Dupont aujourd'hui?").	(1.a) Focus on key words in the stimulus sentence (1.b) Focus on elements of L2 grammar and phonology in the stimulus sentence
2. **Identification and selection without retention:** Respond physically to given commands ("Mettez l'object jaune sous l'objet rouge"). [2 to 4 hearings, as preferred]	• Focus on key words in the stimulus sentence
3. **Identification and guided selection with short-term retention:** Listen to conversations [between 1 and 4 times, as preferred] and state who is talking to whom, and where (e.g. woman ordering meal at restaurant). This involved larger units of speech than tasks 1 and 2 (which were based on single sentences).	1. Focus first on key words and/or key phrases/key sentences. 2. These enable subjects to infer the context and identify roles. *Also*: 1st sentence or 1st question-response sets up expectations on what might logically occur within the context of the conversation.
4. **Identification and selection with long term retention:** Listen to one passage [between 2 and 4 times, as preferred] and later recall as much as possible. [It appears that the passage in question had a title: "Portrait de Daniel(le)", that subjects could -and did- use as an advance organiser]. Topic: a college student introduces herself, her likes and dislikes (incl. school subjects). [highly familiar context].	1. Attempt to develop an overall understanding of the auditory material. 2. Fill in details or particulars that were not initially understood. 3. Attempt to organize the auditory material for recall. [Does not specify whether each stage relates to a new hearing of the input]
5. **Anticipation of elements of sequence:** Listen to a passage in which every 7th word has been deleted and supply the missing words. First played 2 (or 4 times) without responding, then cloze-listening with tape stopped at end of each sentence for [written?/oral?] response.	1. Attempt to determine the context by focusing on those segments immediately before or after the missing word. 2. Choose a word that appears to fulfill all the demands placed on the ommitted word.

Table 4-5: Task-related differences in strategy use (summarised from Defilippis 1980)

Merlet and Gaonac'h (1995) conducted the only experiment in which playback behaviour is explicitly examined in relation to specific listening strategies. Using a CD-ROM in EFL as input and 72 French undergraduate students as subjects, the researchers recorded the listeners' keystrokes to examine their playback strategies under each of the conditions summarised below:

Task steps:	*Experimental conditions:*	**(A)**	**(B)**	**(C)**	**(D)**	**(E)**
First hearing	Straight through, synchronised with a sequence of still pictures	Same for all groups				
Between 1st and 2nd hearings	Advance organiser, presented as a written intro/summary in L1	Yes	Yes	No	No	No
Second hearing (allowed to pause)	Still pictures shown in synchrony with relevant utterances	Yes	Yes	Yes	No	No
	Free to repeat any segment(s)	Yes	No	Yes	No	Yes
After 2nd hearing: ***Recall in L1***, with cues presented in the original sequence	Cues used = full pictures? • *Yes:* same pictures as 1st time • *No:* numbered speech bubbles, labelled with speaker's name	Yes	Yes	Yes	No	No

Table 4-6: Experimental set-up in Merlet and Gaonac'h (1995)

The aim of this experiment was to examine the effects of advance organisers, still pictures and the option to repeat the input on recall. Each idea-unit found in the students' recalls was allocated to one of three categories [my English translations]:

- ideas related to the *propositional meaning* of the text ("informations de type 'analytique' ")
- *integrative elaborations*, e.g. summarising, adding information, inferencing ("informations de type 'global' ")
- *wrong guesses* ("informations de type 'devinement' ")

The presence of pictures during the second hearing was found to generate more pauses and shorter repeated segments. There were also more "wrong guess" ideas in the recalls produced in this condition. Those who were allowed repetition made significant use of their ability to replay segments, and recalled more ideas overall, but they did not pause more frequently than the other groups. In the "pictures" condition, listeners who could repeat produced significantly more "wrong guess" ideas than did any group when no pictures were present. Results related to the use of advance organisers are not reported here, since it would appear that the texts that were used gave not only contextual information, but also global propositions from the dialogue itself (not too surprisingly, these ideas were found later in the recalls).

The greater number of pauses made in the "pictures" condition is interpreted as an effect of greater processing load due to the additional information conveyed by the pictures. This seems plausible, especially as the pictures were line drawings that required a high degree of semiotic decoding (speech bubbles containing visual representations of the meanings of spoken utterances). The higher number of "wrong guesses" (that the authors interpret as a sign of text-driven strategies) seems more related to the presence of pictures than to repetition itself: repetition did increase the number of wrong guesses in the pictures condition, but not in the other conditions.

The studies reviewed in this section show that, in any listening task that uses recorded input, playback procedure and strategy use are indeed closely related. These studies include a few initial attempts to determine what specific relations may operate between the two, as Merlet and Gaonac'h suggest:

> The frequency of interruptions, span of backward moves and frequency of repetitions provide indications as to the level of processing that is being employed (Merlet and Gaonac'h 1995:218, my translation).

This particular study opens a promising perspective. Unfortunately, the analysis of playback behaviour is complicated by other variables introduced in the study, such as the use of advance organisers and pictures.

The patchy findings reported in this section reveal the need for a more systematic focus on the relations between playback behaviour and listening strategies. Once a number of specific playback-related behaviours have been explicitly integrated within the framework of listening strategies, much needed observational data should become available for listening strategy research.

4.2.3 Effective and less effective listeners

A substantial proportion of strategy research is concerned with identifying the strategies that "effective"/"successful"/"skillful" listeners use in comparison to those of "ineffective"/"unsuccessful"/"unskillful" listeners (1992b; Bacon 1992c; Defilippis 1980; O'Malley, Chamot, and Kupper 1989; 1990, study 3; Vann and Abraham 1990). In other cases, the comparison is between different proficiency levels (e.g. using students enrolled on courses at different stages of the curriculum) (O'Malley and Chamot 1990; Vandergrift 1996; Vogely 1995). A common assumption in this kind of study is that successful listeners owe their success to the correct use of effective strategies (Laviosa 1990; 1993; Vandergrift 1996; Vogely 1995). Therefore in order to

become successful, all that unsuccessful listeners need to do is learn how to use appropriate strategies themselves. Another interpretation, however, is that successful listeners use those effective strategies because they have reached a stage of development that allows them to do so (low-level processing has become sufficiently automatic to free up some memory capacity for other types of processing), whereas unsuccessful listeners have simply not reached that stage (Just and Carpenter 1992; Rabinowitz and Chi 1987; Zwaan and Brown 1996).

In two studies using introspection during listening tasks in ESL, O'Malley and Chamot (1989; 1990, study 3) found significant differences between the strategies used by effective and ineffective listeners. Effective listeners were found to use more self-monitoring, elaboration (relating new information to prior knowledge), and inferencing strategies than ineffective listeners. In terms of Anderson's (1985) three-stage model of listening comprehension, the perceptual stage elicited more attentional strategies on the part of effective listeners (see also Vogely 1995), whereas ineffective listeners were more easily distracted (e.g. by the occurrence of an unknown word). At the parsing stage, effective listeners used a combination of top-down and bottom-up strategies, listening for larger chunks, making inferences, and only focusing on individual words when comprehension broke down. Ineffective listeners, on the contrary, focused primarily on bottom-up processing, devoting most of their efforts to decoding the meanings of individual words. Finally, at the utilisation stage, effective listeners used more elaboration strategies than less effective listeners. When exposed to recordings that were particularly long or difficult, listeners (especially less effective ones) appeared to focus on translating earlier portions of the passage and miss out subsequent portions.

Vann and Abraham (1990) conducted a case-study of two female students from Saudi Arabia who had been identified as unsuccessful learners of English. The authors concluded that unsuccessful learners use just as many strategies as successful learners. When their strategies are tallied against one of the strategy taxonomies in use (e.g. Rubin 1981), the results produced by both groups are not significantly different from each other. Defilippis (1980) compared the listening strategies of 13 skillful and 13 unskillful listeners (US college students of French), and also found that the range and overall number of strategies used in the two groups were "more similar than dissimilar" from each other. The differences were in the ranking of each strategy's total frequency within each of the groups. The two lists below show the strategy rankings for each of the two groups (in decreasing order of

frequency. Strategies marked with an asterisk indicate strategies that were reported more often by the group in question (two asterisks = reported more than twice as often as the other group):

> **Skilled listeners:** *using key-words, *contextual inferencing, *using grammar knowlegde, **automatic flow (allowing stimulus to flow automatically without translating), *role identification (identifying speakers' status), **using cognates, **visualization.
>
> **Unskilled listeners:** **translation, using key-words, contextual inferencing, using grammar.
>
> (Adapted from Defilippis 1980)

This comparative ranking shows that, while both groups make extensive use of key word strategies, the most striking difference between the two is that unskilled listeners systematically attempt to apply translation strategies whereas skilled listeners do not.

Stage	Response	Gloss
Noise (N)	None.	No response made because input was heard as "noise"
Distraction (D)	None.	Processing is blocked by a split of attention between the aural input and the previous item.
Syllable restructuring (S)	Restructuring (e.g. tape says "I seem...", subject hears the syllable "ice" and selects the picture of an ice cream)	Listener can hear only a part of a key word and restructures it into a part of a word that was not in the utterance. Association is then based on the misheard key word.
Syllable identification (S)	Projection of a syllable to a key word (e.g. tape says "the washing of cars", subject hears "wash" and selects the picture of a car wash).	Listener finds a plausible word or syllable that could be associated with an icon. The syllable or word is projected onto that word.
Key word association (K)	Key word (e.g. tape says "an artist might use this for painting", subject hears "artist" and selects the picture of a palette).	Listener hears a single key word in the input and matches it to a single icon.
Linked key words (L)	More than one key word is heard (e.g. tape says "Before the development of air travel, this was the only way to travel overseas", subject hears "air"+"overseas" and selects the icon of an airplane).	Listener establishes a pragmatic link between the two key words and associates them with an icon.
Phrases (P)	Whole phrase is processed (e.g. subject hears only the beginning of an utterance "This is an automobile which people hire..." and selects the icon of a taxi).	As whole phrases are processed, function words can serve to link key words contained in the phrase.
Complete images (C)	Entire utterance is comprehended (e.g. subject hears every word the tape says in that utterance).	The utterance is chunked into short-term memory and retrieved as the listener scans the field of icons.

Table 4-7: Picture selection strategies (adapted from Ross 1997)

Ross (1997) examined the levels of processing (syllable, word, phrase, complete utterance, etc.) and inferencing strategies used by 40 Japanese learners of English in a listening task where they had to match 25 utterances to their related pictures from a set of 45. He identified eight different levels of processing of the aural stimulus ("stages"), which were largely related to proficiency (shown on Table 4-7, p.94).

More proficient subjects performed higher levels of processing than low proficiency subjects, and this generally (though not always) led to successful performance. Ross's discussion is consistent with capacity theories:

> ...more proficient listeners are capable of integrating more elements from the aural stimuli into a referential schema that leads to a viable instantiation to a visual referrent [=finding a picture to match it]. The more a listener can extract clues from the input and juggle them in STM, the larger the number of candidate referrents she can potentially instantiate them to. (Ross 1997)

The study also showed that subjects at all proficiency levels made extensive use of key word processing, which appeared to be the major source of input for inferencing strategies.

Vandergrift (1998) conducted a case study in which he used talk-aloud techniques to examine the strategies used by 7 learners of French at different proficiency levels (2 beginners, 3 lower-intermediate, 2 upper-intermediate). He concluded that metacognitive strategies are of key importance for successful listening comprehension. He also suggested that within metacognitive strategies, *monitoring* would appear to be "a superordinate strategy, directing other metacognitive strategies such as prediction and selective attention in addition to cognitive strategies such as inferencing and elaboration" (Vandergrift 1998:392).

The studies reviewed so far tend to converge on two main differences in strategy use between skilled and unskilled listeners: skilled listeners make better use of metacognitive strategies and rely on both top-down and bottom-up processing, whereas unskilled listeners rely almost entirely on laborious bottom-up processing and fail to use metacognitive strategies effectively. A closer look at these differences suggests that they could be related to differences in processing capacity (Just and Carpenter 1992; Zwaan and Brown 1996). Rabinowitz and Chi propose an analysis of strategy use that accounts for individual differences in terms of capacity theories. In a study of the reading strategies used in L1 by children with learning disabilities

(Rabinowitz and Chi 1987), they distinguish two approaches to strategy use, that they call *general-context strategies* and *specific-context strategies:*

- *General context strategies* are chosen at the start of a task, on the basis of the existing constraints and previous experience with similar tasks;
- *Specific-context strategies* are adopted in response to specific task features that are identified while performing the task in question.

The authors claim that, although general-context strategies are easily learnt through training, the use of those strategies alone does not improve performance significantly. Their argument is that general-context and specific-context strategies require different types of knowledge: general-context strategies are related to metacognitive knowledge. For instance when asked to memorise a word list, a subject may initially decide to look for semantic categories within the list (even though he does not know yet whether or not the list will include any possible categories in which to group the words). On the other hand, the subject might only decide to use that strategy once the list has been presented, because she notices that there are several names of animals in it. This second case would be an instance of context-specific strategy. Yet the specific decision can only take place if the subject is able to notice the implicit presence of the category "animal" within the list, a type of response that requires easy access to relevant conceptual knowledge (the word "cat" must activate the node "animal" in the first place). The authors insist that it is not so much a matter of lacking the required knowledge (most subjects would know that "cat" belongs to the "animal" category!), but rather failing to activate the relevant nodes (not linking the word "cat" to the "animal" category when it comes up in the word list). In line with capacity theories of language comprehension discussed in section 2.2.1.1 (Just and Carpenter 1992), Rabinowitz and Chi argue that the necessary connections may well be already in place, but they are just not strong enough to operate at the existing levels of activation. An additional factor could be the absence of a direct connection between the L2 word and the concepts related to it. The need to take a longer route via the L1 (where the only existing connection is located), would also prevent low-proficiency listeners from activating the required connection when processing resources are under pressure (Dufour and Kroll 1995).

An interesting aspect of the theory proposed by Rabinowitz and Chi (1987) is that it accounts for the complex interaction between conscious/intentional strategies and unconscious/automatic processes:

> Any given concept can initially become activated through a variety of goal-oriented (strategic) or nongoal-oriented (spread of activation) processes. Once the concept becomes active, however, it will spread activation to its neighbors regardless of the goal-oriented behavior going on. Thus, the spread of activation affects the activation levels of concepts in a manner that is independent of any specific goal-oriented or strategic behavior (Rabinowitz and Chi 1987:86).

The implication of this theory is that specific-context strategies will only be accessible to certain individuals: those whose cognitive networks have connections strong enough to activate the nodes that the strategy requires in order to operate efficiently, and who also have processing capacity to spare. If one assumes this interpretation, metacognitive strategy training should be of little help to less proficient listeners. Instead, these subjects would stand a much better chance of success in the long run by means of intensive practice at cognitive level.

4.3 Strategy training for listening skills

A logical question deriving from the previous discussion is whether there is any point in training ineffective listeners in strategy use. This section reviews research findings on this matter. Section 4.3.2 describes a number of training appproaches that have been proposed for developing strategy use in L2 listeners.

4.3.1 Can strategy training improve performance?

O'Malley and his colleagues (O'Malley et al. 1985b) conducted a study in which 75 high school students enrolled in ESL classes received strategy instruction for 50 minutes daily over a period of eight days. Training focused on a listening task (5 minute lecture on an academic topic) and a speaking task (2 minute presentation on a familiar topic). The strategies taught for listening comprised one metacognitive strategy (selective attention), one cognitive (note taking), and one socioaffective (cooperation). For speaking, they comprised one metacognitive strategy (functional planning), and one socioaffective (cooperation). In addition to a pre-test and a post-test, daily tests were also conducted to monitor progress. Cues to use the taught

strategies were provided initially for scaffolding purposes, and then gradually removed throughout the 8-day period.

Although post-test analyses for the speaking test were significant in the predicted direction beyond the .01 level, the post-test analyses for the listening test approached, but failed to reach significance. However, analyses of the daily tests did show that the treatment groups significantly outperformed the control group. The authors explain that the scaffolding cues may have been removed too quickly (8 days is not a very long period), and that "the strategy training may have failed to compensate for the complexity of the listening task" that was used as a post-test. Their overall conclusion is that strategy instruction "can facilitate learning", though they also admit that "students presented with a listening task that is too difficult may derive little help from learning strategies" (O'Malley et al. 1985a:576-77).

Thompson and Rubin conducted a classroom-based longitudinal study of the impact of strategy training on Russian listening comprehension over two semesters (Rubin and Thompson 1992; Thompson and Rubin 1993; 1996). Thirty-six US undergraduates taking 3rd year Russian were assigned respectively to the control and experimental groups. The latter received training based on the following types of video-recorded input: scripted material simulating authentic input, drama (clips from feature films), TV interviews and news reports.

In addition to general cognitive strategies (predicting content, listening to the known, listening for redundancies, to intonational clues and resourcing), students were taught cognitive strategies that were more specific to each genre: for drama, focus on the storyline; for interviews, attention to the question-and-answer sequence; for news items, consider *who, what, where, when* and *how*. Among the metacognitive strategies presented to the students for training purposes (listed below) a considerable proportion was closely related to playback strategies. In other words, those strategies implied decisions involving the use of playback equipment to control exposure to the input. They are shown in italics on the list below (my emphasis):

1. Planning, e.g. *deciding how many times to view a particular segment, whether to view it with the sound on or off, determining how to break up the segment into manageable portions.*
2. Defining goals, e.g. *deciding what exactly to listen for*, determining how much needs to be understood.
3. Monitoring, e.g. assessing one's comprehension, identifying sources of difficulty, *isolating problematic portions.*

4. Evaluating, e.g. assessing the effectiveness of strategies used.

(Thompson and Rubin 1996: 335)

The results of the video-based post-test show that students who received strategy instruction improved significantly over those who did not. Improvements of 10% or more were recorded in twice as many cases for the experimental group than for the control group for this test. The authors report less improvement with the more demanding segments (those in which vision and text were not redundant, e.g. interview), and argue that this could be due to a proficiency threshold for strategy training for this kind of listening task. Overall, it is concluded that systematic strategy training does improve listening performance. It must be pointed out that, throughout the study, the control group had only two uninterrupted exposures to each recording, whereas the experimental group had (and used) the option to listen again. Also, the purpose of the listening task was different in both groups, regardless of training factors: the control group used the input "as a basis for speaking and writing activities" whereas the experimental group was explicitly focusing on listening. The results might have been different if the control group had been given the opportunity to engage in tasks in which the primary output related to listening comprehension itself.

In another study involving high-school learners of Spanish, Rubin and her colleagues (Rubin, Quinn, and Enos 1988), trained subjects in the use of three cognitive strategies (prediction/verification, cognates and story line) for listening comprehension based on video recordings. They reported that the groups that had received training outperformed the control group when presented with challenging tasks, but that no significant differences in performance were observed when the tasks were easy. This apparent contradiction between the results of both studies suggest that the effectiveness of strategy training might operate only within the limits of a threshold and a ceiling of proficiency. Outside these, the task may just be too easy or too difficult for effective strategy use.

4.3.2 Approaches to strategy training for listening skills

According to Vann and Abraham (1990), what poor performers lack is the metacognitive knowledge to choose strategies that are suitable to the task in hand and to apply these appropriately. Rabinowitz and Chi (1987) point out that attempted strategy use may fail for two possible reasons: subjects may be using strategies

deficiently (by failing to choose an appropriate strategy for the task in hand), or inefficiently (by using an appropriate strategy in the wrong way). In order to overcome these problems, a variety of strategy training approaches have been proposed in recent years. An ideal training scheme should enable learners to decide when to use a particular strategy, to know how to use it effectively, to remember to keep using it once the training is over, and to transfer it to new listening tasks as appropriate. Three issues in particular have been the focus of discussion:

1. How explicitly should strategies be taught?
2. What role should strategy training play in the L2 listening syllabus?
3. Which particular strategies should L2 listeners be taught?

The first issue relates to the nature and amount of information that learners ought to be given about the strategies they are being taught. "The more the better" seems to be the answer according to a review conducted by Brown and Palincsar (1982) in which they report a number of strategy training studies aimed at children with learning difficulties. The authors identify three possible training methods:

- *Blind training* (often used with very young children): the relevant strategies are demonstrated to the learners, who then practise them. No further information is given about their benefits or rationale.
- *Informed training:* in addition to demonstration and practice, the learners receive instruction in the significance of each of the strategies presented. Each strategy is also given a name.
- *Self-control training:* In addition to all the above, learners receive explicit instruction in planning, checking and monitoring the activity themselves (i.e. applying metacognitive strategies to the actual strategies being taught - e.g. "is it working?").

The findings reported suggest that the third approach is the one that gives children (especially older ones) the best opportunity to develop strategies that are both durable and transferable. Brown and Palincsar conclude that

> ...an ideal training package would consist of both practice in the use of task-appropriate strategies, instruction concerning the significance of those activities, and instruction concerning the monitoring and control of strategy use. (Brown and Palincsar 1982: 7)

In a later experiment reported by Thompson and Rubin (1996) Rubin and her colleagues applied the same three approaches to the training of three groups of high-school learners of Spanish as a foreign language. A set of three cognitive strategies (prediction/verification, cognates and storyline) was taught to each of the groups using the "blind", "informed" and "self-control" approaches respectively. Although the improvements in performance found in all three experimental groups were significantly greater than in the two control groups who had received no strategy training at all, differences in improvement between the three training conditions failed to reach significance (Rubin, Quinn, and Enos 1988, quoted in Thompson and Rubin, 1996). The promising results reported by Brown and Palincsar (1982) have nevertheless encouraged a number of learner training programme designers (Dickinson 1987; Ellis and Sinclair 1989; Fernández-Toro 1999; Fernández-Toro and Jones 2001) to adopt a self-control approach to learner training. Throughout the eight skill-specific sections of their study guide, Fernández-Toro and Jones (2001) present a series of needs-specific tasks (or "DIY techniques"), each covering the same three headings:

1. What is it good for?
2. How to proceed
3. How to assess progress

This ensures that all three areas recommended by Brown and Palincsar (1982) are explicitly covered: (1) learners are informed of the benefits of each particular strategy; (2) they are then guided step-by-step through the relevant procedure; and finally (3) they are given means of evaluating the strategy's effectiveness for themselves. Since the three steps are presented in a variety of forms depending on the specific task to which they apply, this approach is also consistent with Wenden's recommended *knowledge-based* approach (Wenden 1995), which is described later in this section.

A second issue in strategy training relates to the status of strategy instruction within the listening syllabus. Mendelsohn (1995; 1998) proposes an approach that is entirely *strategy-based* . It comprises three basic phases:

1. Making students aware of how the language works ("metalinguistic awareness")
2. Making them aware of the strategies that they already use ("metastrategic awareness")
3. Training them in the use of additional strategies

This is to be achieved through a strategy-based syllabus, for which Mendelsohn proposes the following list of units as a "prototype":

- Strategies to determine setting
- Strategies to determine interpersonal relations
- Strategies to determine mood
- Strategies to determine topic
- Strategies to determine the essence of the meaning of an utterance
- Strategies to form hypotheses, predictions, and inferences
- Strategies to determine the main idea of a passage

(Mendelsohn 1995: 138)

Each of these units is designed to focus a different range of strategies to cover each of the following stages: awareness and consciousness-raising, prelistening activities, focusing the listening, guided activities, practice with real data, doing something with what has been comprehended.

Wenden (1995) prefers what she calls a *knowledge-based* approach to learner training, in which "knowledge about the nature and purpose of the task" is the central concern. She argues that...

> Learner training syllabi and lesson plans should not be strategy driven - derived from a list of strategies [...]. Rather [...], learners should learn to plan, monitor and evaluate learning related to a particular task [...]. The cognitive strategies they are taught should also be specific to that task.
>
> (Wenden 1995: 190)

Her proposed approach focuses on three basic steps:

1. Task purpose (Why should I do the task?)
2. Task classification
   (What kind of task is this? Have I done/learned something like this before?)
3. Task demands (How should I do the task?)

The tasks to be included in the training programme, or *target tasks* (e.g. buying a train ticket, taking lecture notes) are defined according to the tasks that the learner actually needs to perform in real life. Each of these *target tasks* is then divided into its component subskills (e.g. techniques for identifying relevant linguistic cues, recording and organising information, etc.) to which she refers as *pedagogical tasks.*

The last, and probably most important issue in strategy training is to determine which strategies lend themselves to formal instruction and which ones do not. A previous

discussion showed that certain strategies (e.g. metacognitive strategies such as planning) cannot operate unless the subjects have enough active connections and processing capacity to support them (Rabinowitz and Chi 1987, p.95 above). A number of researchers regard strategy training primarily as a matter of promoting the transfer of effective L1 strategies into the L2 (Goh 1998; Mendelsohn 1995; 1998). They claim that ineffective L2 listeners are simply experiencing difficulty in transferring their existing L1 strategies into the L2, and that their problems can be corrected by means of strategy training. An alternative position, known as the *linguistic threshold hypothesis* (Bernhardt and Kamil 1995), argues that such transfer cannot take place unless a minimum level of proficiency has first been reached in the L2.

Zwaan and Brown's findings (1996, see discussion p.46) are consistent with this hypothesis. While they do show that effective L1 readers find it difficult to apply their usual L1 strategies to the L2, they also indicate that the solution may not be a simple matter of strategy transfer. In their study, even proficient L1 readers were not able to integrate information across sentences when reading in the L2, and focused on text-based processes instead. The reason for this appeared to be a lack of efficient lexical and syntactic processes in the L2. No amount of explicit strategy training is likely to counteract this kind of handicap when input demands are simply beyond the processing abilities of the listener. In such cases, practice in bottom-up processing may be more appropriate.

An increasing number of researchers postulate the importance of training L2 listeners in bottom-up processing. Field (1997) stresses that students need to be trained in dealing with assimilation and variation phenomena when they listen to continuous speech, and that such skills "need to be taught, not just practiced" (Field 1998). Cauldwell (1996) proposes the use of computer software that allows learners to observe the different sound shapes that words can take in fast speech. He argues that this type of training can aid perception and comprehension of fast speech in an L2. Another original approach is proposed by Gilbert (1995), who uses pronunciation exercises in order to train bottom-up processing skills for listening purposes. Her argument is that speech production is the best means of making features such as rhythm and intonation explicit to the learners. She even (rather bravely) suggests the use of kazoos in the language classroom to focus attention on those features without the distraction of verbal content.

In a controversial review of learner training studies, Rees-Miller (Rees-Miller 1993) pointed out a number of problems associated with strategy instruction:

- Research findings on the effects of strategy training cannot be conclusive, because strategies "are defined so broadly that it is questionable whether they can be specified in terms of observable, specific, universal behaviors".
- The strategies taught may be "based on cultural models that are not universal" (e.g. silent rote learning vs. vocal self-expression).
- The notion of "learning to learn" can be perceived as patronising by adult and experienced learners, who have already developed their own, tried-and-tested techniques for learning.
- Learner training would tend to assume ideal conditions (e.g. "a selective group of highly motivated individuals with homogeneous goals"). Unfortunately this is not always the case.
- There is often a conflict between the learners' beliefs about how to learn a language and those of the teachers, which "may result in the students' continued clandestine use of techniques of which the teacher disapproves".
- Some strategies may be only suitable for subjects with a particular cognitive style. It is essential to adopt a flexible approach that allows individual choice.

Effective language learners have been reported to become "somewhat impatient" when they are subjected to large amounts of explicit strategy instruction (Chamot 1995). Over-training may also take the form of too many strategies presented at once, resulting in student confusion (Chamot and Kupper 1989).

The issues listed above ought to be considered carefully before any recommend-ations can be made regarding strategy instruction.

4.4 Methods of investigation in strategy research

4.4.1 Common methods used in strategy research

A variety of research methods are used in order to identify the listening strategies that L2 learners use. One of them consists of strategy questionnaires such as the MASQ (Metacognitive Awareness Strategy Questionnaire), which was developed by Carrell

(1989) to examine reading strategies, and adapted for listening by Vogely (1995). This particular example takes the form of a 1-5 Likert scale covering areas such as the subject's perceived ability in the relevant skill ("I am able to..."); repair strategies ("If I don't understand something, I..."); strategies perceived as "effective" ("In order to read/listen effectively, I..."); and aspects of the skill that make it difficult. Oxford (1990) developed a similar questionnaire, called the Strategy Inventory for Language Learning (SILL), designed in a format that is suitable for guided self-evaluation as well as for data collection purposes. It comprises 6 sections, covering: memory strategies, cognitive, compensatory, metacognitive, affective and social strategies. Unlike MASQ, SILL does not focus especially on a specific skill, since it is intended to function as a diagnostic instrument for general L2 learning skills.

Other researchers use talk-aloud protocols in which subjects report their thoughts while performing a specific task (Vann and Abraham 1990). With listening tasks, it is common practice to insert pauses in the recording, during which listeners can verbalise their thoughts, sometimes in response to probe questions such as "What are you thinking?" or "What did you understand?" (O'Malley, Chamot, and Kupper 1989; O'Malley and Chamot 1990, study 3; Vandergrift 1998), but in other cases the subjects have full control of the tape and verbalise whenever they see fit (Long 1991).

Interviews in which learners are prompted to reflect on the strategies that they use are another method for collecting strategy data. Bacon (1992a; 1992b; 1992c) used retrospective interviews, conducted immediately after the listening tasks under study. While listening, subjects were told to make a sign to the researcher "whenever they caught themselves thinking about what they heard, then to focus on what they were doing in trying to understand". This was followed up immediately after by an interview in which listeners reported on their strategies, comprehension, recall, background knowledge, perceived success and affective response.

Sometimes the interviews focus on general listening strategies without referring to a specific task performed beforehand (Vandergrift 1996). Other interviews cover a whole range of learning experiences and skills inside and outside the classroom, as did two of the studies reported by O'Malley and Chamot (O'Malley and Chamot 1990, studies 1 and 2). In this case the subjects were interviewed together in groups of two to five, either using their L1, or the L2 for more advanced levels.

Learner diaries can also provide useful strategy data. Goh (1997) used the method to examine the listening strategies of forty 19 year-old ESL learners in the People's Republic of China. Her printed template prompted introspective reports about specific occasions in which the subjects listened to English, what they did in order to understand better, their thoughts about learning to listen, and how they practised their listening after class. In this study, the diaries were also used longitudinally for strategy training purposes.

Finally, other studies focus on the actual output of a specific listening task in order to analyse the type of strategies that the subjects have used. In particular, mistakes made by listeners convey valuable information for the researcher. Error analysis has proved especially useful in the study of inferencing strategies in listening comprehension (Kelly 1991; Ross 1997).

4.4.2 Introspective studies: dangers and precautions

Introspection is commonly used for data collection in strategy research (Bacon 1992a; 1992b; 1992c; Carrell 1989; Goh 1997; Long 1991; O'Malley, Chamot, and Kupper 1989; O'Malley and Chamot 1990, study 3; Oxford 1990; Vandergrift 1996; Vandergrift 1998; Vann and Abraham 1990; Vogely 1995). The method however has some detractors (Nisbett and Wilson 1977; Seliger 1983), who argue that it is not possible for subjects to consciously give accurate reports of unconscious mental processes, and describe introspections as "conscious verbalizations of what we think we know" (Seliger 1983:183). Matsumoto (1994) reviews some of the most typical concerns about the validity and reliablility of introspective verbal reports. The central issue is whether one can consciously access information related to strategic processing. According to certain authors, this is simply not possible (Nisbett and Wilson 1977; Seliger 1983). Matsumoto argues that it is, so long as verbalisation is concomitant with the task in question, calling upon Ericsson and Simons's argument whereby processes involving information that is currently held in STM are accessible for introspection (Ericsson and Simon 1984).

On the basis of Matsumoto's arguments (1994) four key principles can be drawn up in order to ensure validity and reliability when using verbal reports:

1. The time interval between processing and reporting should be reduced as much as possible. Reports given while information is still in STM are best.

2. Subjects should not be required to follow complicated verbalisation procedures. Reporting must not interfere with performing the task as naturally as possible
3. One should elicit information directly related to a specific task, rather than information that subjects could infer from previous learning experiences.
4. Triangulation is essential. Wherever possible introspective data ought to be interpreted along with observational data.

In line with these basic principles, the present study uses a range of data collection methods which are discussed in detail in Chapter 6.

All the arguments discussed above apply only to controlled processes. Automatised processes are not accessible to conscious introspection, even while the relevant information is still held in STM (Ericsson and Simon 1980; Ericsson and Simon 1987).

Here again, the argument about strategies and consciousness is of central importance. If one adopts a strict definition of strategies as conscious actions (Rabinowitz and Chi 1987), the matter is no longer an issue, given that strategy research will, by definition, not be interested in unconscious mental processes. Unfortunately, a previous discussion (p.79 above) already showed that such a clear-cut separation of conscious strategies and unconscious mental processes is far from easy to establish. For that reason, the present study focuses on overt behaviour, regardless of the role of consciousness in such behaviour. However, it does attempt to operationalise the notions of controlled versus automatic processing on the basis of available data (section 6.2.2.2 below). In line with Seliger's recommendation (Seliger 1983) introspection is only used at the initial stages of the study (Experiment 1), for hypothesis-generating purposes. In subsequent experiments, an observational approach is adopted for hypothesis-testing.

4.4.3 Classifying strategies

Another problem faced by strategy researchers is the lack of consistency across existing strategy taxonomies, and the difficulty of establishing objective classification criteria to standardise the analysis of introspective data across the field (Rees-Miller 1993). While the main categories (cognitive, metacognitive and social/affective) are broadly accepted, virtually every researcher proposes a different list of strategies based on the classification criteria that they regard as most relevant (Bacon 1992a; 1992b; 1992c; Defilippis 1980; Goh 1998; Laviosa 1990; 1993; Long 1991; O'Malley,

Chamot, and Kupper 1989; O'Malley and Chamot 1990; Oxford 1990; Rubin and Thompson 1992; Vandergrift 1998; Vann and Abraham 1990).

Even within the internal system adopted by particular authors, the process of analysing data from think-aloud protocols and retrospective interviews can be problematic. Looking at the subjects' responses that are occasionally reported verbatim in some of these studies, one cannot but notice a considerable amount of subjective interpretation in assigning a particular response to a given strategy category. Efforts reported by the researchers to ensure inter-rater reliability do not seem to prevent the occurrence of ambiguous interpretations. Consider the following example:

Tape plays:	"Chaque semaine au restaurant il y a un bon spécial".
Subject:	Something, let me think, 'il y a' something, I don't know.
Interviewer:	No ideas as to what he said?
Subject:	Not really. 'Bon'... He said 'bon'. I'm not sure

(Vandergrift 1998:275).

The researcher's interpretation of this protocol reads: "any attempts to focus on grammatical detail appear to interfere with the perception of semantic clues". One could just as easily argue that "il y a" (as a chunk) is more salient to the subject because it was probably learnt very early on. One could also say that this student, who had identified in a previous section that the recording had something to do with food, picked up the word "bon" because the schema "good food" had been activated, but failed to notice "chaque semaine"or "spécial" because more specific components of the "restaurant" schema had not. Any of these interpretations (and a good few more) could be used as the basis on which to categorise this particular response in terms of strategy use, and yet none of them can be proved more valid than the other on the grounds of the protocol alone.

Bacon (1992b) quotes the following comment from a subject:

> The word that came after "casa" is still bothering me. "Remolque". That's when I get out my dictionary and look it up.
> (Bacon 1992b:328)

and assigns it to a "strategy category" called "motivation", glossing it with the comment: "successful listeners expressed interest in finding out the meaning of the passages or in learning new lexis". However, there is no evidence here to support a direct cause-effect relation between motivation and dictionary lookup. A listener of lower proficiency might have expressed just the same "interest in finding out the

meaning" of the word "remolque"... if only s/he had been able identify the segment in question as a word in the first place! Motivation cannot be claimed to be the determining factor for strategy use on the sole basis of the student's remark.

In the same study, the fact that "successful listeners were both more precise in their summaries, and provided more detail" in their recall protocols is interpreted as another "strategy category" termed by the author as "detail" (Bacon 1992b:327-28). A more plausible interpretation of such detailed recalls would be that these are *effects* of successful listening, rather than instances of strategy use in themselves. The same can be argued about the status of "motivation" (above) as a "strategy category".

These are only a few examples, but many more could be cited from a variety of studies. Consequently, one has to be cautious in assessing claims made on the basis of statistical analysis of protocol data (however significant the resulting figures might turn out to be) when these rely on variables that are not rigorously defined in the first place. Defining a clear-cut strategy taxonomy is not an easy task, but it is possible to limit the scope for subjective interpretation by operationalising the strategies under study into a set of objective, clearly defined criteria (e.g. Ross 1997; Zwaan and Brown 1996).

When investigating "visible" strategies such as communication strategies, researchers can rely on a number of observational methods (Kasper 1997) because the production process generates an observable output that becomes available for analysis. When, on the other hand, a study focuses on listening strategies, one is faced with an additional difficulty, because the process of listening in itself does not entail any externally visible reponse. O'Malley and Chamot (1990:87) note a distinction between "overt" strategies (behaviours observable externally, such as dictionary use) and "covert" strategies (only accessible through introspection). With the exception of communication strategies in interactive listening set-ups, few overt strategies are readily available for direct observation in naturalistic listening situations. The advantage of using recorded input is that the playback patterns used by listeners ("pause", "rewind" and "play") can be regarded as instances of overt behaviour that can then be operationalised for the purpose strategy research. Long's comment that "the use of a tape-recorder as a record-keeping instrument is a very effective research tool" (Long 1991) points in this new direction. It is interesting to note that this remark was found in one of the very few listening strategy studies in which subjects had been given full control over playback.

However, it must be noted that even a detailed analysis of overt playback strategies cannot provide information as to the cognitive, metacognitive and affective strategies underlying the specific playback patterns used. In order to obtain this information, some form of verbal output needs to be collected from the subjects regarding whatever goes on in their minds during playback. It is for this purpose that the present study uses paired listening protocol analysis in Experiment 3.

5. Research questions for this study

As explained earlier (section 1.4), the aim of the present study is to seek an answer to the following questions:

1. What are the specific features of PL?
2. How can PL be used for research on listening processes and strategies?
3. What does initial PL research tell us about L2 learners' listening processes and strategies?
4. How can PL be used in order to develop listening skills in L2?

In order to investigate these questions, three experiments were conducted. The first two are hypothesis-generating, whereas Experiment 3 uses protocol analysis techniques to test the hypotheses generated in experiments 1 and 2. This section discusses the original hypotheses that were formulated in order to answer the general questions above.

5.1 Experiment 1: Pauses

Experiment 1 studies three conditions:

(A) straight playback repetition with no pauses

(B) repetition with pre-set pauses at regular intervals

(C) repetition with pauses inserted at the subject's discretion.

Full details of the procedure are given in section 7.1 (p. 136).

As reported in section 2.2.2, studies have shown that the insertion of pauses is more beneficial to comprehension than slowing down speech rate (Berquist 1994; Blau 1990). Pauses are believed to help L2 listeners because they give them extra time to complete pending processes and thus reduce processing load (Just and Carpenter 1977) before the next segment is played back. Therefore they are likely to find playback with pauses easier to process than uninterrupted playback.

Since individuals vary in their STM capacity, they are also likely to experience different degrees of demand on their processing resources. If each individual is allowed to time the pauses according to his/her own needs, it is suggested that subjects will find processing in this condition easier than in the controlled pauses

condition. For the same reason, low proficiency listeners are likely to experience greater overload and therefore require more pauses than higher proficiency listeners.

Three hypotheses[13] can be formulated on the basis of this argument:

- *Hypothesis a_1:* subjects will find condition B more beneficial than A, and condition C more beneficial than A and B.
- *Hypothesis b_1:* recall scores will be highest under condition C and lowest under condition A.
- *Hypothesis c_1:* under condition C, low-proficiency subjects will make more pauses than high-proficiency subjects.

The study conducted by Blau (1990) also showed what appeared to be a threshold of language proficiency at which pauses were particularly beneficial. If one interprets such differences as the result of variations in processing load between individuals at different proficiency levels, then one should also expect similar variations to be found at different levels of input difficulty. In other words, if, with a particular recording, the load is so high that no segment of the input can be processed successfully, then pauses will be of no use. If, on the other hand, the load is perfectly manageable at the given speed, then no pauses will be required either.

The following hypothesis derives from this discussion:

- *Hypothesis d_1:* under condition C, subjects will make more pauses when the difficulty of the input is neither too high nor too low for their proficiency level

As well as testing the four hypothses above, a qualitative study of the data was undertaken to find any patterns emerging from introspective and observational data:

- Analysis of the subjects' reported perception of pause-related processes and strategies
- Analysis of observed pausing patterns in group C.

[13] For reference clarity across the chapters of this thesis, the hypotheses tested in experiments 1 and 2 will be named with a lower case letter, followed by a subscript number indicating in which of the two experiments the hypothesis was tested. A complete new set of hypotheses will be formulated for Experiment 3, integrating the findings of the first two experiments.

5.2 Experiment 2: Individual moves in free playback

In Experiment 2, a record was made of the operations performed by 16 subjects on their cassette players while listening to two recordings in L2 as part of a class test for course assessment purposes. Responses were in the form of written answers to a set of comprehension questions in L1. Full details of the procedure are given in section 8.2.

5.2.1 Taxonomy of playback patterns

The primary aim of this experiment was to develop a taxonomy of playback patterns and define a number of operational concepts for use in data analysis in this area of strategy research.

In the process of elaborating such a taxonomy, it was also expected that a set of typical playback patterns would be identified for further examination in Experiment 3.

- Data analysis in order to define a suitable taxonomy and terminology for the study of playback strategies
- Data analysis in order to identify recurrent playback patterns for further investigation in Experiment 3.

In addition to these qualitative analyses, a set of preliminary hypotheses was formulated on the basis of previous research and common teaching practice (sections below).

5.2.2 Repetition and recall rate

As discussed in section 3.2.2.2, repetition has been found to have more beneficial effects on listening performance than increased redundancy by means of signalling or paraphrasing (Chaudron 1983; Chiang and Dunkel 1992). It was therefore hypothesised that this phenomenon would be reflected in the subjects' playback strategies in two ways:

- *Hypothesis* a_2*:* within the sample, subjects playing the recordings more times will have better recall scores
- *Hypothesis* b_2*:* within each recording, segments played more times will be better recalled by the subjects

If proved true, these hypotheses were to be explored further in Experiment 3 in order to examine the possible causes of such a phenomenon through protocol analysis.

5.2.3 Playback patterns: the "a-b-a" hypothesis

The core hypothesis of this study was that the playback patterns produced by individuals across the sample would be susceptible to classification into a number of typical patterns.

Since no research was found specifically related to playback strategies, hypotheses at this stage were largely based on intuitive assumptions deriving from common teaching practice. The starting assumption was that typically, the first and last things that subjects do in a listening task is play the recording straight through. Between these two stages, they were assumed to listen to shorter sections, one at a time with repetitions if needed, progressing through the recording in sequential manner. This hypothetical pattern was termed the *a-b-a pattern.* This pattern is based on a reading strategy described by some authors (Anderson 1985; Nuttall 1996:129) as the SQ3R technique[14], which comprises five stages:

1. *Survey* - Skim text to see if it's relevant and to get overview of main points.
2. *Question* - Pause to ask yourself questions you want the text to answer [note the term "pause"].
3. *Read* - Looking for answers to your questions and noting any other relevant information.
4. *Recite* - Process in some way the salient points gained from the text [this is done by talking aloud].
5. *Review* - Relate this to previous knowledge and experience.

(Nuttall 1996:129)

Another instance of a similar pattern in language teaching is the procedure traditionally followed in dictation exercises: first the teacher reads the text straight through, then with pauses during which students write down what they hear, and finally straight through again while students check throught the text they have written. A number of teaching institutions use the same procedure in listening examinations, where the recording is played in sections after a first run through in order to allow time to write down the answers to a series of questions, sequentially distributed along the recording (Brown and Yule 1983b). The final run through is then used by students to check through their answers. Research findings would support this approach to playback in listening tests: Sherman (1997) found that the first time a recording was heard, subjects could either try to construct a global meaning interpretation or make

[14] Anderson uses a slightly different term in the section entitled "How to use this book" of "Cognitive Processes" *(Anderson 1985)*. He refers to the PQ3R technique: preview, question, read, recite, review), but his description is along the same lines as the one proporsed by Nuttal.

sense of the test questions, but not do both at the same time. The condition in which the highest test scores were observed involved a playback procedure that allowed subjects to first focus on the global meaning, then on specific questions. In Sherman's experiment, this was achieved by presenting the test questions only after the first hearing.

The initial hypothesis of Experiment 2 was based on the a-b-a playback procedure, with scope for revision as appropriate in Experiment 3.

- *Hypothesis c_2, or the "a-b-a" pattern hypothesis:* most subjects will follow a three-stage playback strategy: starting by playing the recording once straight through, then working through it in sections, and ending by playing it one last time straight through.

5.3 Free playback in pairs

In Experiment 3, students worked in pairs through a recording, trying to understand as much as possible without taking notes. They were free to operate their machine in any way they wanted while another machine recorded everything they did with the tape and everything they said throughout the task. Four recordings of varying difficulty were used with different pairs. Full details of the procedure are given in section 9.2 (p.183).

Two very general hypotheses were initially formulated, with a view to develop a more specific formulation once the hypothesis-generating stage had been completed:

A. Similarities in playback patterns will be observed across populations performing identical tasks, and such trends will be different for different types of input.

B. Specific relationships can be established between a given set of playback patterns and specific cognitive, metacognitive and social-affective strategies.

A further hypothesis could be tested at a later stage, if A and B were shown to be true: "Given identical tasks, individual variables (e.g. proficiency, learning style, motivation) will affect the playback strategies that subjects use." However, a qualitative analysis of individual differences is beyond the scope of this study, given that subjects were not working individually in Experiment 3.

Hypotheses A and B were to be operationalised once the results from experiments 1 and 2 became available, on the basis of the playback patterns observed in the

hypothesis-generating phase. The final list presenting the specific hypotheses tested in Experiment 3 can be found on page 181.

6. Research instruments for the study of PL

Two aspects of PL require the design of purpose-made research intruments in order to operationalise the specific variables under study: the first aspect is related to meaning. This section considers not only the assessment of comprehension by means of recall protocols, but also ways of studying how meaning is organised in terms of the temporal sequence in which ideas are presented in a given recording (*idea-maps*) and their level of redundancy. The second aspect relates to the specific playback strategies that listeners actually use. This looks at the way in which subjects sequence their exposure to the input by "moving" backwards and forwards through the recording. This "physical" aspect of strategy use is analysed by means of *move logs*. Finally, the instruments discussed below attempt to relate specific playback strategies to certain cognitive, metacognitive and affective strategies. The method used for this purpose is protocol analysis, based on a talk-aloud paired listening task.

The proposed methods are described in this chapter, and will be critically evaluated in chapter 10.

6.1 Operationalising meaning-related variables in ARI

6.1.1 Recall protocols

Given that the questions used for assessing listening comprehension act as advance organisers and could determine specific playback strategies (see section 3.1.2.3), it was decided to use a totally open-ended eliciting procedure in the form of written recalls for Experiment 1. In order to ensure that the subjects who had control over the tape did not use the pauses to take notes (which would have made their condition significantly different from the other two), none of the groups was allowed to take notes while listening. The protocols were to be written in English for three reasons: the learners had a low level of L2 proficiency, this would prevent writing skills from interfering with listening skills, and subjects would not be able to repeat verbatim L2 segments that they heard unless they understood their meaning.

6.1.2 Idea-units

The term *idea-unit* has often been used for analysing samples of spoken and written language; this term is generally applied to units that are present in the output that

subjects produce (Crookes 1990; Kroll 1977). In the present study, however, idea-units are used for analysing comprehension rather than production.

It would of course be over-simplistic to claim that a particular idea is simply "contained" in a given segment of a recording. It is now widely accepted that comprehension is not just about decoding the propositional content of a message, but a process of interpretation in which the listener is just as active as the speaker in constructing meaning (Brown and Yule 1983b; Rost 1990). Nevertheless, for the purpose of assessing comprehension quantitatively in this study, simple declarative messages such as "the doctor's secretary takes calls on Tuesdays and Thursdays" can safely be regarded as "containing" a number of ideas. For the sake of clarity, the metaphor of segments "containing" ideas will be used hereafter. This is on the assumption that most of the ideas that were used as idea-units in the present study would have been interpreted in very much the same way by any competent speaker listening to the same message. Such assumption was further supported by the high inter-rater reliability scores obtained in the measures using the proposed idea-units.

Meyer (1975:100) used idea-units for analysing recalls of written prose by L1 readers. Her definition of the term *idea-unit* covers both content units and existing relationships between these content units. In a simplified view of her model, ideas are arranged hierarchically, with the highest ones forming the macro-structure (gist) and the lowest ones forming the micro-structure (detail). A similar scoring system is adopted in this study, although relationships between content ideas are spelt out in the form of simple statements. By reducing the terminological conventions required, raters scoring the subjects' recalls do not need to be trained to distinguish theoretical categories at an abstract level. The simplified version used in this study was intended to assure good reliablity while using a straightforward set of categories.

As an example, let us examine a recording in which the speaker says "My name is Miguel". The scoring system must be able to differentiate the various levels of comprehension reflected in the following responses, all of which are typical examples of those found in the subjects' recalls:

Speaker:	"My name is Miguel"	
Recall:	(1)	His name is Miguel
	(2)	Miguel [as an isolated word]
	(3)	I can't remember his name
	(4)	His name is Manuel
	(5)	Manuel [as an isolated word]
	(6)	[No recall]

Table 6-1 shows how the use of 4 different idea-units allows subject 1 to be given a higher score than subject 2 on her recall of the original utterance, while the least accurate response (subject 5) can still be credited above total lack of recall (e.g. subject 6).

Using this technique, an idea-unit checklist was produced for each of the recordings from

Idea No.	***Subjects' responses →*** ***↙ Ideas:***	(1)	(2)	(3)	(4)	(5)	(6)
1	A name is given	✓	✓	✓	✓	✓	
2	The speaker's name is given	✓		✓	✓		
3	Miguel	✓	✓				
4	The speaker's name is Miguel	✓					
	Total scored →	4	2	2	2	1	0

Table 6-1: Idea-unit scoring grid

which recall rates were to be collected. The design of idea-checklists comprised two steps:

First a pilot version was produced, including a series of statements, words and phrases (the proposed idea-units) believed to represent the propositional meaning of the entire recording. This first checklist was tested using the recall protocols produced by the sample.

A second version of the checklist was then produced in order to ensure better discrimination across the sample. If a particular group of related idea-units all produced a discrimination index of 1 (meaning that every single idea-unit in the group had been recalled either by all the subjects in the sample or by none of them), those ideas were collapsed into one global proposition. For instance: if the entire sample had recalled all four ideas listed in Table 6-1 (above), these could have been amalgamated into the single proposition: "The speaker's name is Miguel". The same would apply if nobody in the sample had recalled any of the four ideas.

If, on the contrary, different levels of recall were observed across the sample for a particular group of related idea-units, the original list of idea-units was preserved, and if necessary, expanded in order to reflect all the different levels of recall that occurred in the sample.

This criterion of discrimination is not present in Meyer's model, given that her aim is to study discourse structure, whereas the aim here is to design a scoring instrument that is

easy to use and an accurate indicator of a subject's relative listening performance within the sample. The drawback is that such idea-units can only be defined post-hoc on the basis of the range of recalls produced by the particular sample that is being studied and the particular recordings that were used as input. This type of idea-unit however, proved to be a highly reliable scoring tool for the intended purpose (see chapter 10 for a critical evaluation).

When the response was cued by specific questions (e.g. recordings E and F, used in Experiment 2), the lists of idea-units were much shorter than those used in listening tasks requiring free recall. This is because the use of questions focuses the subjects' response onto the limited set of information that is relevant to the questions, thus reducing the range of possible responses.

6.1.3 Measuring redundancy in PL

In the specific context of PL terminology, *redundancy* refers to the number of times a particular idea is either explicitly mentioned or directly inferable from the recording. The most reliable way to assess whether a particular segment of a recording has been understood is to assess the recall of those ideas that are mentioned only at that particular point in the recording. Such ideas mentioned only once in a recording will be referred to as *non-redundant ideas* (NRIs).

Ideas that are mentioned more than once, or can be inferred from more than one segment in the recording will be referred to as *redundant ideas* (RIs). Successful recall of RIs is more difficult to relate to a particular segment, since RIs are present in more than one place.

In addition to its original level of redundancy within a particular recording, a given idea will have its redundancy level raised each time the relevant section of the recording is played again. This feature, which is specific to PL, implies that any tool used for assessing comprehension in PL must take into account both measures of redundancy:

1. The level of redundancy of each idea-unit within the recording.
2. The total level of redundancy reached by each idea-unit as a result of repeated playback.

Because Experiment 1 was only concerned with the effectiveness of pauses in assisting comprehension, only two levels of redundancy were considered (NRI = mentioned once; RI = mentioned more than once). If the same instrument is to be used in subsequent

studies, a more refined measure is likely to be needed for RIs. Such a measure could take the form of a *redundancy index*, which would indicate the precise number of times the idea-unit was actually mentioned, either in the original recording or as a result of repeated playback. The use of idea-maps (described below) allows for the integration of redundancy indices if needed.

6.1.4 Idea-maps

As the two main operations available to subjects (pausing and repetition) affect the timing of input exposure, the first consideration is time-related.

Reading research traditionally uses a range of techniques for the study of time-related strategies. These include time-controlled input by means of a computer that displays the input in successive chunks, response-time measurements based on items of varying length and nature, pages that are turned over only after completing a given task in order to see the next passage, and eye movement observation in unrestricted conditions.

In listening, the temporal aspect becomes even more relevant due to the demands of on-line processing (see section 2.2). In addition to this, PL allows subjects to alter the sequence of exposure to the input depending on the playback procedure adopted. Any system designed to study PL must integrate this sequential dimension as well as the semantic dimension represented by the location of idea-units in a given recording.

6.1.4.1 Segments

All the recordings were divided sequentially into segments of more or less similar duration. Such sequential units (hereafter called *segments*) normally coincide with utterances, but can sometimes be longer or shorter than the utterance depending on speech rate. The boundaries between sequential units are arbitrarily defined by the researcher, their length depending on where it would be possible to introduce a pause in the recording without breaking the natural flow of speech[15]. This definition of segments responds to purely mechanical, rather than theoretical, criteria, and represents only one of the many ways in which a recording could be divided into sequential units. As this study is not primarily concerned with discourse analysis, this criterion for sequential division seemed adequate for the purpose of representing the temporal dimension of recorded speech.

[15] Blau (1990) used a similar criterion for pause insertion in her study on the effects of pauses on listening comprehension.

These tape segments - which are defined post-hoc by the researcher - are not related in any way to the process of segmentation that is carried out by the listener as a cognitive step in word recognition. In parts of the discussion where ambiguity could be a problem, the terms "tape segment" and "speech segment" will be used respectively.

6.1.4.2 Idea-maps

Using idea-units and segments, an *idea distribution map* (or *idea-map* for short) can be drawn up for each recording. Idea-units represent the vertical axis and segments the horizontal axis (Table 6-2, p.124). In this way, both the semantic and sequential/temporal dimensions can be represented at once

The top row shows the numbers corresponding to successive segments of the recording, in sequential order. The shaded row just below indicates the speakers taking part in the recorded conversation. In this case, A (the researcher) is asking questions (this is represented with a question mark: "A?") and B (the Mexican student) is replying to the questions.

The different idea-units are represented as numbers[16] under the segments of the recording in which each is mentioned. Shaded boxes indicate ideas that are only mentioned once in the recording (non-redundant ideas); whereas numbers in standard type indicate ideas that are mentioned more than once. In that way, the idea-units "contained" in the portion of text played back by the subject during a given move can be easily identified.

The bottom line shows the main sections in the text (these are simply to give a rough indication of what the idea-unit numbers refer to).

For Experiment 1 (in which one group of subjects was allowed to play the recording only once, but could interrupt it with the "pause" button at any point), pauses made by a given subject can be represented as vertical lines on the idea-maps, thus allowing to see at which point of the recording they were made, and which ideas were present in the segment that immediately preceded each pause. It would also be possible to measure and record pause duration under each of the vertical pause lines for analysis purposes.

[16] The full list of idea-units corresponding to these numbers can be found in Vol.2, Appendix 4.4. See also Vol.2, Appendix 2.4 for a complete transcript and translation of the recording.

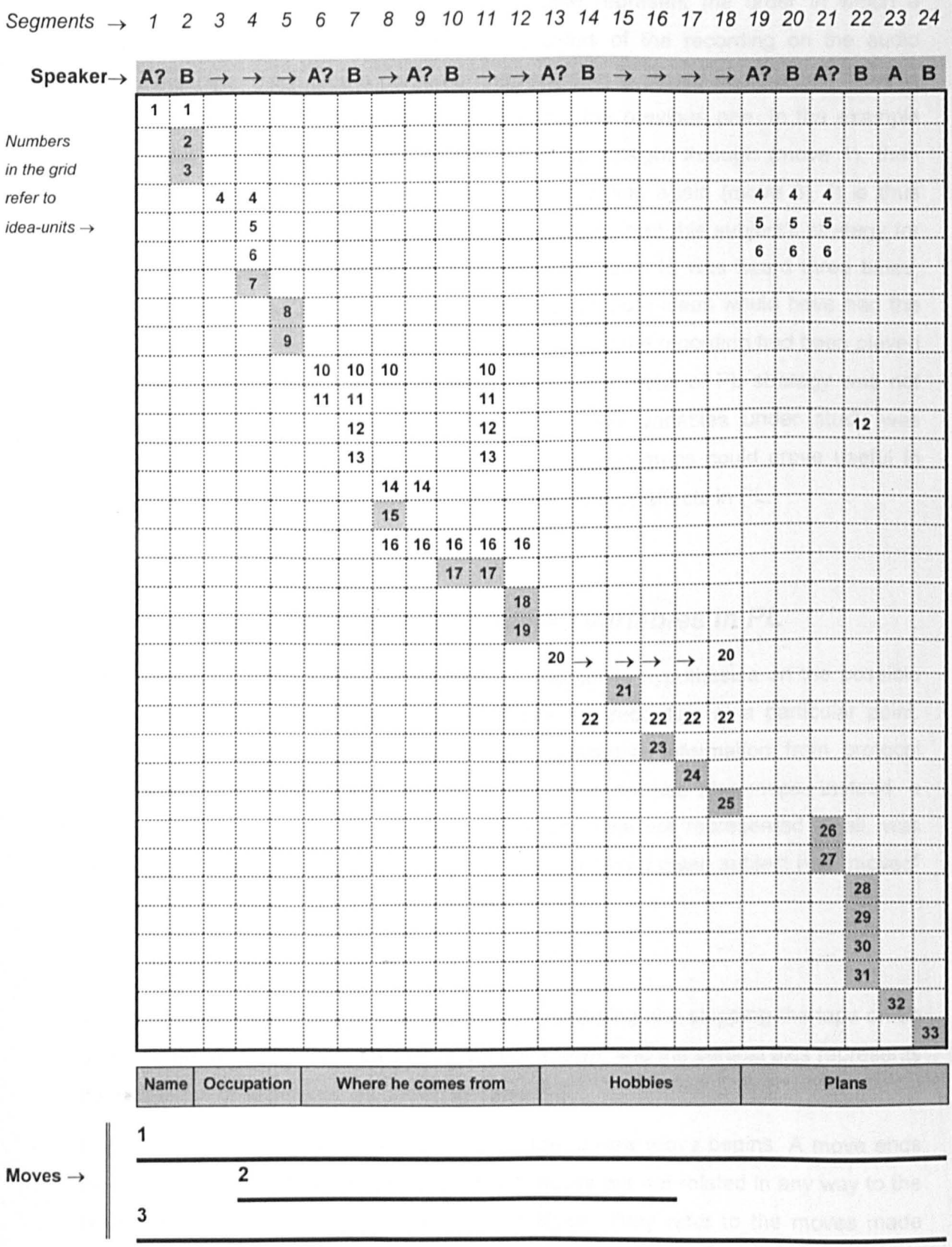


Table 6-2: Idea distribution map for recording A2.

The *moves* recorded at the bottom of the table represent the order in which a particular subject plays back the different sections of the recording on the audio playing machine. Each time a portion of the recording is played back, a new move is added in the form of a new horizontal line under the previous one. In the example shown here, the subject first played the recording straight through (move 1), then segments 4 to 13 (move 2), finally the whole recording again (move 3). It is thus possible to establish the level of redundancy resulting from this subject's strategy for each of the ideas present in the text. For instance, idea 15 was heard three times, whereas idea 32 was heard only twice, even though both ideas would have had the same level of redundancy (both were originally NRIs) if the recording had been played only once. A systematic analysis of redundancy as a result of PL strategy was not conducted in this thesis, since the number of new variables under study was considerable already. However, it is believed that idea-maps could prove useful in future studies that focus more specifically on redundancy effects in PL.

6.2 Operationalising strategy-related variables in PL

Idea-maps were only used in Experiment 1, in order to hypothesise on the possible factors that could have made a subject pause the recording at a particular point. Since it was hoped to obtain more reliable qualitative information from protocol analysis in Experiment 3, experiments 2 and 3 did not use idea-maps. Instead, a simpler system, in which the semantic dimension was not represented at all, was used: *move logs* that show exclusively the way in which a given subject has "moved" through the recording during a listening task.

6.2.1 Move logs

The term *move* refers here to operations such as playing and stopping the tape on an audio player. In a move log, idea-units are not shown, and the vertical axis represents the sequence of segments, as shown in Table 6-3.

Every time the subject presses the "play" button, a new move begins. A move ends as soon as the subject stops the tape. These moves are not related in any way to the term "move" as it is used in conversation analysis. They refer to the moves made when operating a playback machine.

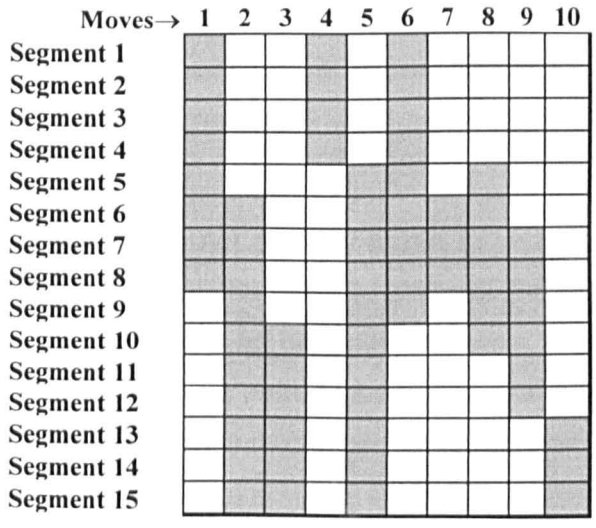


Table 6-3: An example of move log

On a move log, moves are represented by shaded columns. The segments of the recording that are played back together in one go are highlighted in the relevant column. This method was used in experiments 2 and 3 in order to describe, categorise and analyse the playback patterns produced by the subjects. See section 8.3.4 for a full set of related terms and their proposed definitions.

6.2.2 Talk-aloud protocols from paired listening tasks

To elicit information about the cognitive, metacognitive and affective strategies underlying specific playback strategies in Experiment 3, some form of process-related verbal output was required from the subjects. The method chosen was to record (with subjects' consent) unsupervised peer discussions during a paired listening task. Output from the playback machine was recorded at the same time, so that the subjects' comments could be related to the relevant moves on their move log.

6.2.2.1 Rationale

Following a previous discussion of introspective methods for strategy research (section 4.4 above), four necessary precautions to be taken in collecting verbal reports were identified:

1. The time interval between processing and reporting should be reduced as much as possible. Reports given while information is still in STM are best.
2. Subjects should not be required to follow complicated verbalisation procedures. Reporting must not prevent from performing the task as naturally as possible
3. One should elicit information directly related to a specific task, rather than information that subjects could infer from previous learning experiences.
4. Triangulation is essential. Wherever possible introspective data ought to be interpreted along with observational data.

The proposed method complies with all these recommendations:

1. Subjects provide verbal information at the same time as they perform the task. Talking is actually the medium through which they negotiate meaning with their task partner.
2. Subjects are not explicity asked to verbalise their thoughts. This occurs naturally as a result of the need to perform the task jointly with a partner.The nature of verbal output is no different from what it would be in a non-experimental learning activity involving paired listening in a language laboratory. The experimental setup is remarkably similar to the natural context in which PL is most likely to be used by L2 learners.
3. Since the only reason for talking is to perform a set task, all the information obtained is directly related to the task in question. Subjects are not required to make judgements as to which processes may be taking place. Verbal information is intended for the task partner, rather than the researcher.
4. Observational information on playback behaviour is collected in parallel to verbal data by means of move logs. Furthermore, the method used for protocol analysis adopts wherever possible an observational approach to data processing.

In spite of these significant advantages, the method also has a two weaknesses:

1. Individual variables are difficult to study, given that the subjects are not working individually.
2. The physical absence of the researcher prevents unwanted intrusion, but it also prevents her from prompting relevant information. The data obtained may fail to explain why certain behaviours were adopted, while conventional interview methods can elicit richer data.

The first problem is difficult to avoid. Some amount of control could be gained by pairing up the subjects according to one variable (e.g. proficiency score). However, Experiment 3 showed that other individual variables such as assertiveness and motivation may prove difficult to control when the subjects are working in pairs. As for the second problem, additional information could be obtained by conducting follow-up interviews: the tape with the pair's conversation could be played back to the subjects as a cue, prompting retrospective comments separately from each partner. This follow-up technique was judged to be beyond the scope of the present thesis, given the sheer volume of data collected from paired listening protocols alone.

6.2.2.2 Conventions for protocol analysis

For the analysis of protocol data, subject's spoken reponses were interpreted as instances of three possible categories: cognitive processes[17], metacognitive strategies, and social/affective strategies. The cognitive category was subdivided into decoding and inferencing processes, following the criteria explained below. For practical reasons, the actual encoding was performed by hand using coloured highlighters, making it difficult to include a full set of encoded protocols in this thesis. However, the conventional symbols shown below in printed form were used on one of the protocols (protocol no.1), which is shown in Vol.2 (p.59) as an example. The conventional categories represented by handwritten and printed symbols are identical in both systems.

6.2.2.2.1 Responses interpreted as decoding processes

Two response types were interpreted as instances of decoding processes:

1. Where chunks of input were uttered by a subject in the L2;
2. Where the wording that was uttered in the L1 matched exactly (or was an attempted match of) the original wording in L2.

Inevitably, some type 2 responses could be mistaken for local inferences (see next section), and vice-versa, particularly in cases where the "attempted" match resulted in inaccurate translation. However an absolutely precise distinction between decoding and inferencing processes at propositional level was not believed to be critical in an experiment of this nature.

Decoding processes were classified according to the type of speech segment to which they were found to apply. Three types of speech segment were considereed:

1. A single word, or a two-word phrase such as "article+noun" (but not "noun+adjective").
2. A phrase: any lexical item longer than a single word (e.g. "noun+adjective") which does not constitute a full tone group (see below).

[17] The distinction between *process* and *strategy* has already been discussed in chapter 4. Since the issue of consciousness/intentionality is irrelevant in the present encoding system (and for concision's sake), the term *process* will be used hereafter for all cognitive categories.

3. A tone group, or any unit larger than a tone group. For convenience, such units were grouped together under the term "chunk".

The smallest unit to be considered as a chunk was the tone-group, defined by Halliday (1989:53) as "a meaningful segment of the discourse" which "contains a particular point of prominence" and a melodic contour in one or two parts. Units longer than a tone group - such as certain utterances and turns (Crookes 1990) - were also encoded as a chunks.

Propositions that represented a synthesis of several propositional units (e.g. "The doctor proposes two solutions") were treated as inferences rather than decoded chunks.

Symbol	*Label*	*Response description*
[L2w]	L2 word	Word uttered in L2 ("amigo")
[L2p]	L2 phrase	Phrase uttered in L2 ("mes y medio")
[L2c]	L2 chunk	Chunk uttered in L2 ("¿Eres de Madrid?"; "me gusta nadar, pero no me gusta correr")
[L1w]	L1 word	Word uttered in L1 ("friend")
[L1p]	L1 phrase	Phrase uttered in L1 ("a month and a half", "my parents")
[L1c]	L1 chunk	Chunk uttered in L1 ("Are you from Madrid?"; "I like swimming but I don't like running")
[L2...]//	No L1 follow-up	An L2 item for which no L1 meaning is ever supplied in the protocol.
[+]	Grouped items	Several items previously decoded are grouped together (where there is no evidence of inferencing - i.e. all information is explicitly stated in the related segment of the message)

The overt behaviours described above can be related to cognitive processes according to the following set of operational criteria:

- Any processing that does not result in an overt verbal response is assumed to have taken place covertly, either as a result of automatic processing, or with so little control that verbalisation was not needed or possible.

- If an item is uttered overtly in L2 (e.g. *"casa"*), all intervening processes between perception and word recognition are assumed to have taken place covertly. Lexical access is not assumed until proved otherwise.
- If an item is uttered overtly in L1 (e.g. *"house"*), lexical access is assumed to have taken place. If this item has not been uttered in L2 before, all intervening processes between perception and lexical access are assumed to have taken place covertly.
- The size of an item when it is uttered for the first time is proportional to the amount of covert processing that has already taken place. In other words: when a full chunk (e.g.*"one of my friend's graduation party"*) is uttered straight away, more processing has already taken place covertly than when the subject utters a single word (e.g. *"friend"*).

6.2.2.2.2 Responses interpreted as inferencing processes/strategies

An inference was assumed to have taken place whenever the L1 response contained information that was not explicitly stated in the original L2 message. The criterion for encoding inferences as either "local" or "global" is based on the definitions used by Rost (1990) which are discussed in section 2.1.3.3. It must be borne in mind that, whereas certain cases were obvious instances of local or global inferencing, a perfectly clear-cut distinction between the two was difficult to achieve in a number of other cases.

Symbol	*Label*	*Response description*
[inf]	Local inference	An interpretation of the lexical/propositional meaning of the segment ("llamó a la policía"[18] → "the police are there").
[INF]	Global inference	An interpretation of the base meaning of the recording ("policía"→ "he mentions the police and earlier he talked about glass on the floor"→"he's been burgled"), or of its pragmatic context ("When he tells her the cost she sounds horrified").

[18] Translation: "She called the police".

6.2.2.2.3 Responses interpreted as metacognitive strategies

Three basic categories are regarded as metacognitive in this study:

1. Global recap (summarising the information gained so far)
2. Strategies related to making and testing hypotheses
3. Explicit accounts of playback strategies

Global recap

The first category is traditionally included as "summarising" or "synthetising" under the cognitive heading in strategy taxonomies (O'Malley and Chamot 1990; Oxford 1990). However, it was thought that monitoring performance was also an important element in this strategy. In fact, most of the time, the subjects' output did not consist of summaries (which imply selecting and grading information within a hierarchy of importance). Instead, it took the form of comprehensive accounts of everything that had been understood so far. For this reason, global recaps were regarded as metacognitive strategies that informed subsequent task-planning, and were closely related to self-monitoring.

Strategies related to making and testing hypotheses

Another metacognitive strategy is hypothesis-testing. It is often triggered by subjects identifying an interpretation as problematic.

- *Identifying a problem*

Almost on every move made, the subjects asked questions that identified certain information items as problems requiring further information or processing. These could be related to any level of processing: auditory perception (e.g. "...and something else. I didn't know what it was"), segmentation/word recognition ("Did he say 'calle'?"), lexical access (" 'celebrando'... what does that mean?"), local inferences ("I don't know what 'sudar' is ... maybe it's hair dryers"), global inferences (e.g."He's in Madrid, so I think Barcelona are coming to Madrid") or gaps found when reviewing the information gained so far ("What was his name?").

Two behaviours are associated with the "problem identification" strategy. The first one is to ask the partner, the second is to play again the segments of the tape containing the problematic input. This may or may not be followed-up by one of the strategies listed below.

A frequent problem when encoding this strategy was that the subjects almost always formulate every hypothesis in the form of a question (see "Social-affective strategies" below), and it is often difficult to establish when the question is genuine and when it is simply a common hedging strategy. For this reason, questions that were immediately followed by further elaboration by the subject him/herself were not counted. The questions regarded as genuinely strategic were those that were followed by an explicit response from the partner, further playback of the relevant segments, or rejection/ confirmation in subsequent moves.

- *Rejecting a hypothesis*

Two behaviours may show that a previously formulated hypothesis has just been rejected: either the subject explicitly makes a negative statement ("Tablets? –No."), or a new hypothesis is formulated in contradiction with the previous one ("carne" > "meat", followed by "Carme... It's a name!").

- *Confirming a hypothesis*

When a hypothesis is confirmed, the subject generally repeats it as an affirmative statement, which may take the form of a single word for low-level hypotheses related to decoding ("veinticuatro"), an exclamation ("Yes!") or a more elaborate statement ("...and then they started to talk about breakfast didn't they"). A frequent case in this experimental set-up is that one of the partners formulates the new hypothesis and the tape is played again so that the other partner gets a chance to check it ("Trabajando. Did you hear?" > plays segment again).

It has to be pointed out that not all hypotheses that are confirmed are necessarily confirmed in an observable manner. Some of them may well be confirmed by the subjects through an internal process that is not explicitly verbalised. This could be the case for some of the information items that were identified as problems by the subjects, but did not appear to generate further responses.

- *Refining a hypothesis*

Some hypotheses may not be rejected or confirmed completely, but modified instead. In some cases, further decoding of related units allows a more accurate hypothesis to be formulated. At a local level, this was often observed to be the result of processing grammatical morphemes such as tense, gender or number markers, possessives, demonstratives and similar units once the core lexical item had been decoded ("Something in the afternoon" > "this afternoon").

Explicit accounts of playback strategies

Playback strategies are strategies where the subjects operate the machine in order to control the rate at which they exposed themselves to the input. They are closely related to deliberate task planning, and were therefore encoded as metacognitive strategies. The following behaviours were used as criteria: "play back segment", "rewind to start", "stop/pause", and "play on".

Protocol data were only encoded for playback strategies in cases where subjects *explicitly* stated their intention to operate the machine in a particular way, or instructed their partner to do so. The remaining moves were recorded on the pairs' move logs, but not encoded as playback strategies on the verbal protocols.

Summary: notation conventions for metacognitive strategies

Symbol	*Label*	*Response description*
[++]	Global recap	A recap of all information gained so far.
[?]	Problem identified	Question or "don't know" statement when followed by explicit action
[!]	Hypothesis confirmed	New statement confirms hypothesis, or solves previously indentified problem
[x]	Hypothesis rejected	New statement rejects or contradicts hypothesis
[&]	Hypothesis refined	New statement adds information to hypothesis without rejecting it
[<]	Play back segment	Verbal instruction/statement of intention to play back the last segments heard
[<<]	Rewind to start	Verbal instruction/statement of intention to rewind the tape right to the beginning of the recording
[Stop]	Stop / Pause	Verbal instruction/statement of intention to stop the tape.
[Play]	Play on	Verbal instruction/statement of intention to play on.

6.2.2.2.4 Other notation conventions used for cognitive and metacognitive processes/strategies

A few additional conventions were used in order to facilitate data analysis:

Symbol	*Label*	*Response description*
[...](=)	Correct	The subject's interpretation is correct[19]. This symbol was normally left out, unless a "correct" response is being compared to an "incorrect" response to the same item.
[...] (x)	Incorrect	The subject's interpretation is incorrect
[...] (~)	Partly correct	The subject's interpretation is partly correct, but with some degree of inaccuracy.
(...)	Repeated response	When a given response has already been observed in a previous move, it is shown in italics, and round brackets are used instead of square brackets.
[...]>	Followed up later	When a response to a given item of input is followed up by another response to the same item later on (e.g. "amigo"→"friend"), this is indicated by a ">" sign after the code (see below).
>[...]	Following from previous	The response that follows up a previous one is indicated by a ">" sign before the code. For instance: "amigo"→"friend" is encoded **[L2w]>[L1w]** if both occur in the same move. If they occur in different moves, first use **[L2w]>** for "amigo", and later **>[L1w]** for "friend".
[...]/[...]	Multiple interpretations	Where the same response can be interpreted in more than one way (either because of ambiguity or because several processes are clearly taking place at the same time), all processes are listed together, separated by slashes.

6.2.2.2.5 Social/affective strategies

The conventions for encoding this group of strategies were somewhat looser than those used for other types of behaviours. Initially, any verbal behaviour susceptible to be interpreted as a social or affective strategy was highlighted in blue on the protocol transcript. Nonverbal audible behaviours such as laughter, smiles detected in

[19] The categories "correct" and "incorrect" are understood respectively as "acceptable understanding" and "misunderstanding", as defined by Rost (1990:62).

speakers' voices, sighing (and other, less decorous noises!) were also noted. The relevant behaviours were examined under a series of possible, often inter-related categories:

- *Reactions to the input:* complaints about speech rate, auditory conditions, empathy/lack of empathy with the speaker or situation presented.
- *Evaluations* of one's own performance, the partner's, or the pair's (both positive or negative).
- *Reactions to the task:* attitude to the researcher, to the experimental set-up, evaluations of the task (easy or difficult, useful or pointless).
- *Attitudes to risk-taking:* use of laughter after formulating hypotheses, hedging strategies, formulating hypotheses as questions, use of deliberately far-fetched inferences for effect.
- *Pair dynamics:* giving or requesting specific information, re-focusing partner's attention onto the task / distracting partner from task, negotiating procedure together, helping one's partner (playing back a segment again, justifying partner's problems, praising the partner or the pair). Refusing to cooperate.

Given the wealth of social/affective information that this method of data collection elicits, only those behaviours that were of special relevance to playback strategies were examined in some detail.

Paired listening protocols appear to have considerable potential for further research on social/affective strategies in a variety of areas. This method integrates a highly effective combination of subject freedom (no researcher present, interlocutor is a peer, presence of recording equipment justified by the nature of the task) and researcher control (every move and every word recorded for later analysis, task variables can be modified as required, recording can be used later as cue for introspective analysis).

This concludes the initial description and justification of the methods used in the experiments that will be reported in the next three chapters. The methodological discussion will be resumed in Chapter 10, where the methods proposed above will be critically evaluated.

7. Experiment 1 - Pauses

The first experiment examined three conditions:

(A) straight repetition with no pauses

(B) repetition with pre-set pauses at regular intervals

(C) repetition with pauses inserted at the subject's discretion.

The following hypotheses were to be tested in this experiment[20]:

- *Hypothesis* a_1 - Subjects will find condition C more beneficial than A and B, and condition B more beneficial than A.
- *Hypothesis* b_1 - Recall scores will be highest under condition C and lowest under condition A.
- *Hypothesis* c_1 - Under condition C, low-proficiency subjects will make more pauses than high-proficiency subjects
- *Hypothesis* d_1 - Under condition C, subjects will make more pauses when the difficulty of the input is neither too high nor too low for their proficiency level.

As well as testing the four hypotheses above, a qualitative analysis of the data was undertaken to identify any emerging patterns in relation to:

- The subjects' reported perception of pause-related processes and strategies
- Observed pausing patterns in group C

7.1 Procedure

7.1.1 Data collection:

- *Subjects:*

The subjects were 39 undergraduate students at the universities of Newcastle-upon-Tyne and Sunderland, taking courses in Spanish as a foreign language at Beginner or Elementary level in 1993-94. All of them had received less than 150 hours' formal instruction in the foreign language.

[20] See section 5.1 (p. 111) for a full discussion of the rationale for these hypotheses.

- *Procedure*:

Pre-test: A proficiency test was conducted, after which the subjects were placed at random` into three groups of similar overall proficiency for the experiment. In the proficiency test, each of three recordings (A1, B1, C1, transcribed in Appendices 2.1, 2.2 and 2.3 respectively) was played once, after which the subjects were prompted to write down in English as much information as they could remember. They were then allowed to proceed to the next recording. Rewinding and note-taking were not allowed. Written recalls were then analysed and scored according to the number of ideas recalled from a set checklist of 186 items (see Appendices 4.1 to 4.3). The scoring procedure was blind-tested for reliability by a rater trained in the use of idea-unit checklists on 10 randomly selected scripts, and gave a highly significant inter-rater correlation of 0.99 (p=.000).

Experiment: A second set of three recordings (A2, B2, C2, transcribed in Appendices 2.4, 2.5 and 2.6 respectively) was played to all three experimental groups, who were asked to produce written recalls after each recording was played twice. Three different playback procedures were followed:

- *Group A:* 1st time straight through; 2nd time straight through.
- *Group B:* 1st time straight through; 2nd time with fixed pauses inserted in the recordings at regular intervals.
- *Group C:* 1st time straight through; 2nd time with the subjects free to pause when and for as long as they wanted, but not to rewind the tape or take notes.

Hereafter, pauses controlled externally by the researcher will be referred to as *controlled pauses* or *set pauses*, whereas pauses that are controlled by listeners themselves will be referred to as *free pauses.*

The same scoring system as for the proficiency test was used, with a set checklist of ideas comprising 182 items (see Appendices 4.4 to 4.6). All tests were given in a language laboratory and 11 subjects from group C (taking the test at different times) were recorded from the master console while they were playing back the recordings for the second time. The resulting data were analysed to establish the number of pauses made, their duration and their distribution throughout the three recordings.

Questionnaire: The experiment was followed by a questionnaire (Appendices 5.1 to 5.3) asking the subjects about their difficulties and where appropriate, the usefulness

of pauses and their own pausing strategies. All three groups were asked to rank the following procedures according to their perceived usefulness:

- Recordings played twice, straight through without interruption.
- With pauses already inserted in the tape the second time round.
- With control on the pause button the second time round.
- Recordings played three times, straight through without interruption

- *Materials:*

The proficiency test and the experiment materials had identical formats: each comprised three recordings of increasing difficulty in Spanish. The first recording (A1 / A2) was a short semi-scripted interview with a student about basic information on his studies, hobbies and prospects (two different students were used for the proficiency test and the experiment). The second recording (B1 / B2) was an unscripted conversation between two native language teachers simulating a phone-in advice programme on a fictitious radio station (the same speakers who recorded the proficiency test swapped roles and made up a different problem-call for the experiment). This recording was of medium difficulty, but longer than the first and third recordings. Its structure was typical of the phone-in genre: greetings, followed by explanation of the problem by a sufferer; discussion of possible solutions with an expert; sufferer's final reaction to the solutions proposed and closing formulae. The third recording (C1 / C2) was a genuine interview with a native speaker talking spontaneously about his occupation at a relatively high speech rate (two different speakers were used for the proficiency test and the experiment).

Throughout the rest of this chapter, the recordings used in the experiment (A2, B2 and C2) will be refered to as "recording 1", "recording 2" and "recording 3" respectively (i.e. according to the order in which they were presented) for clarity purposes.

- *Data processing:*

An idea distribution map was drawn up for each of the three recordings. This technique is fully explained in chapter 6 (see p.124). Each idea-unit was labelled differently depending on whether it was redundant or non-redundant in the recording (whether it occured several times or only once). Finally, pauses made by each of the subjects recorded in group C were represented as vertical lines across the idea-maps. The researcher was then able to examine which particular idea-units were

present in the segment of the recording that immediately preceded each of the pauses made.

7.2 Results

The data from this experiment were studied from the following perspectives:

1. Perceived usefulness to learners of different playback procedures (from questionnaire data).
2. Learners' accounts of their own pause-related strategies (from questionnaire data).
3. Observed pausing strategies (based on case-study of the pausing patterns recorded in group C and recall protocol data).
4. Effects of pauses on performance (based on recall protocol data)

7.2.1 Perceived usefulness of different playback procedures

Average group preference rankings are represented in Table 7-1 on a scale of 1-4, with 1 being highest preference. The results show that all 3 groups preferred free pauses to fixed ones. There was also general agreement on the fact that, given only two chances to hear the recordings, it was best to have pauses of one kind or another rather than none at all the second time round.

Procedures ↓ Overall preference ranking ↘	Group A	Group B	Group C
Recordings played twice, straight through without interruption.	4	4	4
The second time round, with the pauses already inserted in the tape for you at the end of each sentence.	2	2	3
Having control on the pause button the second time round.	1	1	2
Recordings played three times, straight through without interruption	3	3	1

Table 7-1: Group preferences on playback procedure

The only noticeable disparity was found between group C and the other two conditions. Group C (the only group who did have control over pause location) was the only one who would have preferred one third chance to hear the recording, even if

this meant listening straight through all three times, thus losing the opportunity to have pauses the second time.

All three groups agreed that two playbacks with no pauses at all was the least useful option.

7.2.2 Students' reported strategies

Subjects' questionnaire comments suggest a variety of functions for the pauses made. The comments quoted below are mostly from the two groups who had an opportunity to evaluate the benefit of pauses for themselves during the experiment: group B (with ready-made pauses) and group C (with freedom to pause where they wanted).

- *Pauses used for rehearsing*

Subjects mentally repeat the last few words heard. When asked whether pauses helped or not, they pointed out the importance of this strategy. "Yes, [pauses helped] to catch *all* the words said." (group B, my emphasis). A member of group C who chose not to use pauses explained: "[pauses] would not have helped me as I could not recall the words/*sounds* I'd needed to make sense of the recording" (my emphasis). In other words, he did try to rehearse but failed to recall the sounds heard.

- *Pauses used to reduce processing load*

By interrupting the input, subjects try to avoid overload. This strategy is described in a typical comment: "[Pauses help] because you don't miss part of the next sentence when you think about what's just been said." (group B). A subject from group B (i.e. who had pauses controlled by the researcher) complained that he would have liked to stop the tape himself at other places, for similar reasons: "I hear words I vaguely remember but need time to think so either I miss the next passage or listen on and forget the word.". However, pauses as an opportunity to ease short-term memory load seemed to be perceived as more effective when used on easy recordings. As the level of difficulty increased, perceived effectiveness of pauses seemed to decrease: "[Pauses in recording 1 (easiest)] gave time for what had been said to sink in. Other recordings didn't help as they broke it up and made me realise how little I understand!" (group B). In this case, pauses not only fail to help, but they also generate anxiety as a negative side-effect.

- *Pauses used for inferencing the meaning of unknown words*

One of the subjects from group A, who did not have the benefit of pauses explains that there were some moments where she would have really liked to slow the speakers down: "Particularly after a word I did not understand. Reflection might have helped to put it in context.". Another subject, who did have control of the tape (group C), wrote: "[Pauses] gave you a chance to break the conversation up and [get a] better understanding". However, subjects generally argue that you can only recognise words that you already know in the first place, suggesting that no amount of pausing or inferencing will make up for substantial gaps in lexical knowledge: "[Pauses helped] if I knew the words. Then (...) I was able to pick them out." (group B); "The pauses helped if I knew what had been said" (group C). This may explain why so many subjects did not find pauses useful in more difficult recordings. One student from group C argued that he did not stop the tape very often on the third recording (the hardest of the three) because "it seemed clearer just letting it play." (group C). Another subject, from group B explained that, whereas pre-inserted pauses in recording one (the easiest) did help, pauses in recording three did not, because "the pauses seemed in unnatural breaks in the dialogue." (group B). A comment cited earlier from a subject in the same group expresses a very similar judgment: "[Pauses in recording 1] gave time for what had been said to sink in. Other recordings didn't help as they broke it up and made me realise how little I understand!".

- *Pauses used to recap information and re-structure it for later recall*

This strategy was often used in group C: "I just tried to concentrate on facts which would not stay in my head". Another subject from group C argued that pauses did "not really" help, and gave this explanation: "For recordings 2 and 3 I stopped the tape but because the conversations were so long I lost my train of thought." A similar strategy was reported in several other cases: "[I paused] just when I needed to recollect my thoughts" (group C); "It gave you a chance to absorb the information." (group B); "Helped a bit. Just to recap in my head in English what had been said" (group C). The last example shows that this subject is already trying to verbalise the recall in L1 at this stage. When asked whether she followed any particular pattern for pausing, the same subject replied "Whenever I felt I needed to remember a bit clearly". This particular strategy could be directly related to the type of response that the subjects were expected to produce in this experiment. Although practical as a data collection technique, recall protocols appeared to generate a fair amount of anxiety and may have elicited the translation strategy reported here.

- *Pauses used for resting*

Some pauses are simply used as breaks in the middle of a demanding task. Here are some of the answers given by group C subjects when asked whether they followed any particular pattern for inserting their pauses: "When I felt I couldn't understand for quite a while and was switching off because I couldn't understand. It was simply to refresh myself"; "When I felt I had completely lost the thread of the conversation"; "Whenever I couldn't cope". In this last case, pauses are not seen to improve comprehension of the preceding segments. If anything, they might provide a fresh start for attempting comprehension of subsequent ones. When asked whether pausing helped, another student from the same group replied: "All it did was give a little respite but it didn't help [me] understand any more" (group C). Difficulty of input appears to increase the need for this type of resting breaks. One subject reported that, although he did not stop very often overall, "on the third recording it was more often" (group C). Fatigue build-up within the same recording was also mentioned: "It would have been easier if [pauses] became longer through the dialogue" (group B). Some of the pauses inserted by subjects in group C were very long indeed, as for instance in recording 2, where one of the subjects made pauses of up to 39 seconds. Yet there was no recall of the ideas preceding such extended pauses. This would suggest that the break was not used for processing purposes, but possibly just to rest.

7.2.3 Observed pausing strategies in group C

The actual pausing patterns of group C were examined for triangulation purposes. Given that a number of factual discrepancies were observed between self-reported and observational data, these will be addressed first.

7.2.3.1 Comparing observed and self-reported pausing strategies

In the questionnaire given to group C, the subjects were asked the question "Did you stop the tape very often the second time round?". Originally, this question was not intended to elicit factual information, given that information on the pauses that they actually made was already available from observational data. The question was simply intended to introduce the next question, that elicited the subjects' reasons for pausing or not (hence the vagueness of the term "very often", which would otherwise have been replaced with a more specific formulation such as "how many times").

Nevertheless, as part of routine triangulation, the subjects' responses to this question were compared to their observed behaviours (results shown in Table 7-2).

Subject	Self-reported strategy *Q: Did you pause the tape very often?*	Observed strategy *Rec.1*	*Rec.2*	*Rec.3*	*Total*
101	Never.	0	0	0	**0**
106	No.	0	0	0	**0**
107	No.	1	5	1	**7**
109	No.	0	1	1	**2**
202	Yes. Once or twice.	0	0	0	**0**
208	No. Not really.	0	4	2	**6**
210	Yes. Fairly often. Not in the last one (easier to let it run).	1	6	2	**9**
306	No.	0	5	8	**13**
505	No. Only a few times.	0	2	1	**3**
602	Never.	0	0	0	**0**
604	Yes. Two or three times.	1	3	3	**7**
704	No. Not very often. More often in the last one.	0	3	4	**7**

Table 7-2: A comparison of self-reported and observed pausing strategies in group C

Out of the 12 complete sets of data obtained from group C, only two are unambiguously consistent: subjects 101 and 602, who made no pauses at all and replied "never" to the question. Overall, the number of pauses made by subjects who gave negative answers presented a considerable spread of values, ranging from 0 to 13. Furthermore, the subject who made most pauses in the entire sample (subject 306 → 13 pauses) was among those who gave a negative answer.

Positive answers also covered a range of values (0; 7; and 9). In the case of subject 202 (who made no pauses at all), the positive answer given was clearly inaccurate. It should also be noted that one subject (604), who made 7 pauses gave a positive answer, whereas the two other subjects who made the same number of pauses (107 and 704) gave negative answers.

In summary, considerable inconsistencies were found in the sample's perception of the pausing strategies that they used. In one instance (202), a subject was even found to report clearly misleading information on specific behaviour. Such findings

were thought to deserve further consideration, and will be addressed in the discussion.

7.2.3.2 Number and distribution of pauses made

In recording 1 (the easiest one), only 3 out of the 12 subjects recorded chose to pause. All three did it at the same point of the recording (between segments 12 and 14).

The other two recordings show a common pattern, with pauses becoming gradually more scarce as the recording progresses. In the first half of recording 2, a total of 19 pauses were recorded across the group, whereas only 8 pauses were made in the second half. The same applies to recording 3, in which 22 pauses were made in the first half and only 8 in the second.

As for individual strategies, four subjects did not pause the tape at all in any of the three recordings. Another 8 subjects paused in both recordings 2 and 3. Finally, one subject was only recorded while working through recording 3, but reported having made pauses in all three recordings.

Subjects (code numbers):	*106*	*208*	*704*	*101*	*109*	*210*	*202*	*306*	*107*	*505*	*510*	*604*	*602*
Proficiency percentiles:	**4**	**6**	**19**	**21**	**28**	**34**	**45**	**55**	**68**	**72**	**81**	**91**	**96**
Recording 1 (easy)	—	—	—	—	—	1	—	—	1	—	No data	1	—
Recording 2 (moderate difficulty)	—	4	3	—	1	6	—	5	5	2	No data	3	—
Recording 3 (difficult)	—	2	4	—	1	2	—	8	1	1	9	3	—
Total pauses made	**0**	**5**	**7**	**0**	**2**	**9**	**0**	**13**	**7**	**3**	No data	**7**	**0**

Table 7-3: Pauses made by group C

The correlation index between proficiency scores and the number of pauses made was only 0.38, p=.45. Therefore the existence of a link between proficiency and the number of pauses made could not be proved.

The decision of not making any pauses at all does not appear to relate to proficiency either, as can be seen from the percentiles of proficiency scores (4%, 21%, 45% and 96%) of the four subjects in group C who did not pause the tape. Two main reasons

were given for not pausing: lack of ability to recall what had been said in the first place, and the fact that pauses delay the recall even more and put greater pressure on memory, especially when note-taking is not allowed.

7.2.3.3 Pause type and strategic trends

From a qualitative point of view, two hypothetical categories of pause location were identified: after a salient word, and at structural boundaries (new speaker or topic switch). A number of pauses remained without an obvious explanation, although a few of these could potentially be interpreted either as a result of anxiety (a "stop that tape!" reaction when input rate becomes too fast). This interpretation was applied to pauses that were abnormally close together, leaving between them segments too short to even contain one idea. These pauses are referred to as "panic" pauses. Table 7-4 summarises this hypothetical classification.

Structural	**Non-structural**	
Pauses at structural boundaries: • at topic switches • at the end of a turn (new speaker) • at the end of an utterance (often marked by falling intonation and pause)	**After a salient word** that is... • followed by a natural pause • stressed • easy to recognise (cognate or very familiar to learner).	**For no obvious reason** In the middle of a turn/an utterance. Some of these could be anxiety-generated (panic pauses).

Table 7-4: Types of strategic pauses

As the table shows, pause type was determined on the basis of pause location. However, pause location cannot be taken as an accurate indication of learner strategies. Indeed, not all subjects were equally skilled in operating the machines fast enough to pause exactly where they intended (this problem was mentioned in the questionnaires). In the absence of related introspective data, each evaluative decision from the researcher increases the potential risk of arbitrary interpretation. For this reason, the following discussion should be regarded as a series of hypotheses rather than results.

Of the subjects who paused the tape, most appeared to use a combination of pause-types, regardless of their level of proficiency. One subject paused only after salient words whereas another one made only structural pauses. Rather than using one

strategy throughout all the recordings, subjects generally seem to use a range of criteria for pausing the tape. Idea-map data suggest the following strategies:

- *Follow text structure:*

Pauses related to this strategy are mostly structural pauses, often made just before or just after a question from the interviewer, but also found on less clearly marked topic switches and, at a lower level, utterance boundaries. Subjects using this strategy tend to have better recall of higher level ideas. The strategy is often adopted by proficient subjects, but there are also signs of attention to structure in less proficient subjects. An interesting pattern was observed in recording 1 (the easiest one). All the subjects who paused the tape did so only once, at virtually the same point in the recording (around segments 12-14), clearly dividing it into the same two structural halves: (1) speaker's identity, occupation and birthplace, and (2) speaker's hobbies and plans for the future.

- *Pick up as many key words as possible:*

In this strategy, pauses are mostly non-structural, after salient words. A sign that it is being used could be the presence of isolated words in the recalls. Once the presence of a key word has been detected on the basis of its saliency or redundancy, several strategies are possible, depending on the subject's passive lexicon, the level of risk that s/he is prepared to take and the amount of information available from other segments and background knowledge. The following typical responses were observed:

1. *Proficient subject's strategy* (if the word is recognised): Integrate the word into the global meaning of the text . A successful pause using this strategy cannot be detected unless the salient word is also non-redundant (i.e. can only have been recognised at that particular point in the recording).
2. *Risk-taking strategy* (if the word is unknown): Make up the meaning from available clues as in recording 3, when subject 306 hears "económica" and writes down "he teaches economics". Having failed to perceive the cohesional clue provided by word ending, the subject uses cognate analogy and top-down inference instead.
3. *Low-risk strategy:* Quote the word in L1 (if it was recognised) or transcribe it as heard in L2 (if it was not recognised), as a single, unconnected word.

4. *Face-saving strategy:* Quote the word as heard in L2. making it clear that its function as a key word (if not the meaning) has been understood. The word may be followed by an attempt of translation, often with question mark, as in recording 3, where subject 210 wrote: "—Are you a 'maestro'? —Yes I am a 'maestro' (headmaster?)".
5. *Avoidance strategy:* Leave out the word altogether (see below).

- *Ignore entire sections (low-risk avoidance strategy):*

When a passage becomes too difficult, subjects do not bother to pause the tape at all. This strategy was adopted by all subjects in the last section of recording 3 (segment 37 to the end). Even proficient subjects like 505 (proficiency percentile: 72%) use this avoidance strategy in recording 3. She makes only one pause after segment 36, then lets the tape run for the rest of the recording. She clearly made the pause at the end of what she could understand (20 ideas recalled before the pause, but none at all afterwards).

- *Stop that tape! (anxiety hinders any strategy):*

Any non-structural pause that does not immediately follow a salient word could be due to anxiety, but this is especially clear when a pause is abnormally close to the previous one, leaving a unit too short to be processed. Subject 306 seems to make panic pauses after a similar number of segments in both recordings 2 and 3, possibly due to a fatigue threshold after around 37 segments. Subjects who produced such apparent anxiety-generated pauses tended to show similar anxiety in their response to the questionnaires (704: "Stopped whenever I couldn't cope"; 306: "It makes me realise my vocabulary is terrible!").

7.2.4 Effects of pauses on performance

The comparative study of average recall scores of groups A, B and C did not show any significant differences between the three groups.

	Group A No pauses	Group B Set pauses	Group C Free pauses
Proficiency:			
Mean score	29.3 %	29.3 %	29.9 %
S.D.	5.9	9.0	11.0
Performance:			
Mean score	31.8 %	33.8 %	34.0 %
S.D.	11.1	12.9	13.1

Table 7-5: Proficiency and performance scores of groups A, B and C

Group C mean score was found to be 2.2% better than group A's and 0.2% better than group B's. Group B score was 2% better than group A's. However, none of these differences was statistically significant.

A qualitative study of recall pointed to the existence of five basic patterns of recall that may or may not be related to pause patterns. The nature of the data makes it impossible to establish a causal relation between the insertion of a pause and successful recall of the segments immediately preceding the pause in question. Consquently, the following observations are largely speculative at this exploratory stage.

Five levels of recall were observed in relation to the segments preceding the pauses that subjects made:

1. *Virtually all ideas in the preceding segments were recalled.* For instance in recording 1 (the easiest), where all those who paused between the two sections recalled the first section very well.
2. *Ideas recalled from the preceding segments were only the most redundant ones.* For instance in recording 3, the list of subjects taught by the speaker (that he repeated twice)
3. *Ideas recalled were the most relevant to text structure.* This pattern appeared several times in recording 2, where three subjects recognised a problem-solution structure, but failed to recall which particular solutions were discussed.
4. *Only isolated words were recalled.* These words might have been picked up either because they were very familiar to the subject ("amigos" in recording 2) or because they were particularly salient in the utterance.

5. *No ideas were recalled from the preceding segments.* This was particularly obvious in the final section of recording 3 (a particularly difficult passage[21]). A parallel pattern in this group of segments was that hardly any of the subjects made pauses at all.

7.3 Discussion

7.3.1 Implications of results for the original hypotheses

Regarding the three conditions: (A) straight repetition with no pauses, (B) repetition with pre-set pauses at regular intervals, and (C) repetition with pauses inserted at the subject's discretion, the first hypothesis of this study was:

- *Hypothesis* a_1*:* subjects will find condition C more beneficial than A and B, and condition B more beneficial than A.

This hypothesis was confirmed: all groups found that having control on pause insertion (condition C) was preferable to having no control on it (condition B). Since not everybody experiences overload at the same time, pauses that are inserted by somebody else may interfere with one's own processing needs. Some of the comments from group B reported in section 7.2.2 (p.140) illustrate this problem.

All three groups also generally agreed that two exposures to the recording straight through (condition A) was not as good as two exposures with pauses of some kind the second time round (conditions B and C). Despite a few negative comments, the overall feeling was that pauses do help.

Interestingly, the one group who had control on pause insertion (group C) would have given up pauses altogether in return for a third chance to listen to the whole recording, whereas the other two groups regarded pauses (of either type) as preferable to one more hearing. The reason for this could be that group C was the only one to experience the actual limitations of listener-controlled pauses (i.e. they may be of little help if you are not allowed to repeat the problematic segments).

[21] This section included culture-specific references to the Spanish educational system within a fairly abstract context (see transcript in Appendix 2.6). See also section 2.2.3: "Factors affecting difficulty in a listening task" (p.50).

Three tendencies emerge from the student's responses to open-ended questions on the usefulness of pauses. The last two are contradictory, and some of the subjects' claims conflict with observational data:

1. It is more worthwhile to pause at the beginning of the recordings (this claim is supported by data from the move logs).
2. Pauses are only worth using if you can understand something in the first place. Of the three recordings, pausing was most beneficial in the easiest one (this claim is not consistent with move log data: only three instances of pausing were observed in recording 1, all of them virtually at the same point in the recording).
3. Pauses are something that you use in order to take a rest when it gets too difficult, therefore more pauses were made in recordings 2 and 3 (this claim is consistent with move log data).

Several students also complained that pauses are pointless unless they can be followed by repetition. Having to produce recall protocols may also have caused anxiety, as pauses delay recall even more.

- *Hypothesis b_1:* recall scores will be highest under condition C and lowest under condition A.

This hypothesis was not confirmed. The lack of statistically significant results may be due to various reasons. Pauses delay recall, so the final score may be negatively affected by pause insertion (as several subjects pointed out). More accurate measurements could be made using methods other than whole-text recall protocols (cloze-tests, for instance). Another factor overriding the effects of pauses may have been the effective use of inferencing strategies. Examples of salient words given earlier show how, whereas certain segments may have been better decoded thanks to the pauses, others are more likely to have been understood through inferencing. Finally, not all the subjects in group C chose to make use of pausing strategies. Consequently, the global scores of group C reflect a mixture of performances that may possibly have been affected by pauses, as well as others that could not possibly have been affected (i.e. those of subjects who did not pause).

The lack of conclusive results could also be explained in terms of the location of preset pauses in recording B. Neither Blau (1990) nor this study specified whether the pauses presented to the subjects occurred at clause boundaries or at hesitation points. However, Voss' experiment using naturally produced hesitation pauses (Voss

1979) suggests that pauses at hesitation points have very different effects from pauses at clause boundaries. One could presume that extending hesitation pauses even more by inserting artificial pauses is detrimental, whereas artificially extended clause-boundary pauses are beneficial. Variations in the duration of pre-set pauses could also have significant effects on comprehension.

- *Hypothesis c_1:* under condition C, low-proficiency subjects will make more pauses than high-proficiency subjects

This hypothesis was not confirmed. Four of the subjects (with different levels of proficiency) chose not to make any pauses at all. In line with the tendencies just discussed for hypothesis a_1, two main reasons could be given for this: the lack of ability to recall what was said in the first place, and the fact that pauses delay the recall even more and put greater pressure on memory, especially when note-taking is not allowed.

- *Hypothesis d_1:* under condition C, subjects will make more pauses when the difficulty of the input is neither too high nor too low for their proficiency level

This hypothesis was confirmed by move log data. Very few pauses were made in recording 1 (the easiest), and often no pauses at all were made in the final half of recording 3 (the most difficult material). However, it must be noted that another tendency could have accounted for the latter: pause frequency clearly decreased throughout recordings 2 and 3, yet the end of recording 2 was not particularly difficult. This pattern suggests that another effect (e.g. fatigue build-up) could be operating within each individual recording. Therefore the hypothesis of two phenomena operating in combination should also be considered.

7.3.2 Subjects' reported strategies

As well as testing the four hypotheses above, a qualitative analysis of the questionnaire data was conducted in order to identify the nature of strategies used in relation to pauses. The results described in section 7.2.2 show that a wide range of metacognitive, cognitive and affective strategies were used. These are summarised below, and triangulated with findings from obervational data.

- *Metacognitive strategies*

Evidence of monitoring strategies was found in the form of directed attention for the task in hand (pausing when you feel that you are switching off). Evaluation was also present, as subjects decided to pause when they could not cope, and also made

judgements about their lexical knowledge. Pauses that followed the text structure or occurred after key words could be interpreted as instances of selective attention.

One metacognitive strategy that could not be easily observed in an experiment of this nature was planning. However some of the structural pauses made by group C may have been the result of planning strategies, given that the subjects had already heard the recordings once before being given control of the tape.

- *Cognitive strategies*

Cognitive strategies were even more apparent: mental repetition was used during pauses for word recognition (rehearsal strategy), pauses were used to restructure information in order to recall it (organisation strategy), and at least low levels of inferencing were attempted during pauses in order to work out the meaning of unfamiliar words. Evidence of higher level inferencing was also found in the recalls, but could not conclusively be interpreted as a direct result of pausing.

- *Social-affective strategies*

Affective strategies were observed in the form of panic pauses and possibly avoidance strategies such as letting entire passages run through when they were simply too difficult. Remarks such as "It makes me realise my vocabulary is terrible!" could also be signs of affective responses to monitoring. Social strategies were not applicable in this experimental setup, given that the subjects worked alone.

7.3.3 Hypothetical pausing patterns

Three general tendencies appear to emerge from student's reports and observational data:

1. *Pauses become fewer and farther apart as the recording progresses.*

   Possible reasons could be that information that is not at the beginning is perceived as less important, or because the cumulative effect of incoming information produces saturation. It must be noted that recording 1 did not follow this pattern, possibly because it was extremely easy.

2. *Structural pauses may occur at topic-switches, new turns and occasionally utterance boundaries.*

   This would explain the single pause made by three subjects in the middle of recording 1.

3. *There appears to be a level of text difficulty (neither too easy nor too difficult for the subject) where pauses are regarded as most useful by the learners.*

   This idea can be related to Blau's suggestion (1990) that there is a threshold of proficiency at which pauses are most beneficial. Blau formulates it in terms of the learner's own proficiency level, whereas this study points to the relative difficulty of the recordings used for input and the relative difficulty of segments within the same recording.

7.3.4 Discrepancies between self-reported strategies and observational data

The lack of consistency reported in section 7.2.3.1 needs to be discussed. As explained earlier, the purpose of the prompt question was not to elicit factual information, and therefore its formulation ("Did you pause the tape very often the second time round?") was somewhat ambiguous. The vagueness of the term "very often" may also account for the vagueness of the replies obtained. However, whilst some ambiguity was originally expected, the sheer range of behaviours observed across subjects giving a similar reply was not. Two individual cases caused particular concern. Firstly, the subject who had paused most often in the entire sample was among those who gave a negative answer to the question. Secondly, one subject who had not made any pauses at all gave a positive answer, implying that s/he had paused "very often".

It is fair to argue that one subject's notion of "not very often" may differ from another's. Nevertheless, if an action as overt as pressing the "stop" button a certain number of times is susceptible to undergo such a range of possible interpretations (ranging from ambiguous to simply wrong), the accuracy of introspective reports on less tangible strategies becomes questionable. Indeed, the question asked was vague, but typical prompts for introspection are also in the form of fairly open-ended questions - e.g. "What are you thinking?" (O'Malley and Chamot 1990), "What do you do to listen effectively?" (Vogely 1995) - which would hardly elicit more precise responses than "Did you pause the tape very often?".

Since this experiment was of an exploratory nature, self-reported strategies were regarded as adequate sources of information for speculative purposes. In subsequent experiments however, only observational data were used, both for quantative and

qualitative analysis. Such an observational approach was adopted in order to prevent subjective mediation on the part of the learners (Seliger 1983).

This is not to say that introspective methods should forever be relegated to exploratory studies: they can be powerful means of understanding what subjects *believe* themselves to be doing. Good awareness of learner perceptions and assumptions is indeed necessary in order to develop teaching strategies accordingly (Rees-Miller 1993). However, the primary concern of this thesis is the actual playback behaviour that learners use in specific PL tasks. Subjects' assumptions about PL would be another interesting matter for investigation, but that would be the subject of a different study.

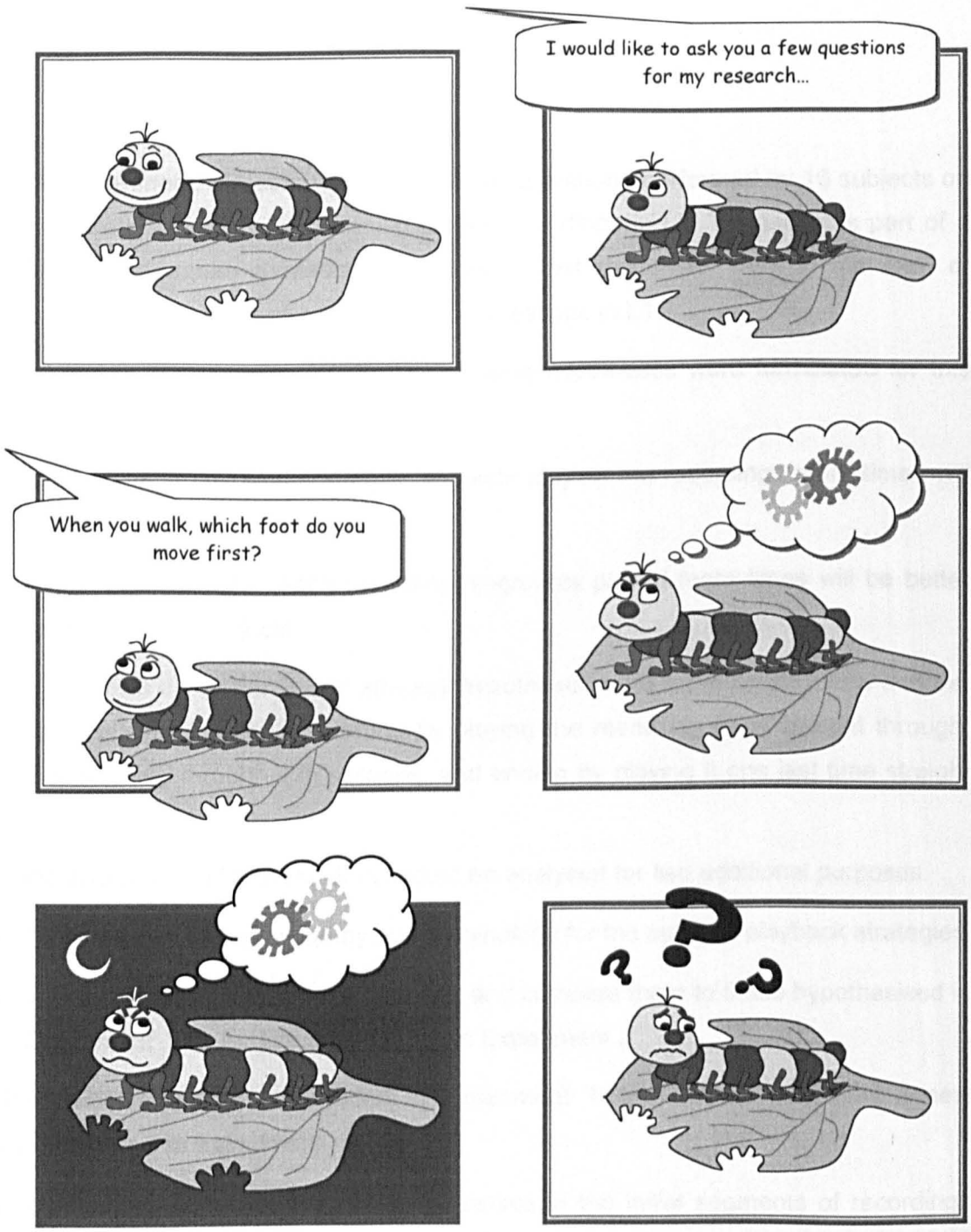


Table 7-6: The introspecting caterpillar (original)

8. Experiment 2 - Individual moves in free playback

8.1 Hypotheses to be tested

In this experiment, a record was taken of the operations performed by 16 subjects on their cassette players while listening to two recordings in L2. The test was part of a class test for course assessment purposes. Test responses were in the form of written answers to a set of comprehension questions in L1.

In an earlier chapter (see p.113) the following hypotheses were formulated for this experiment:

- *Hypothesis* a_2: within the sample, subjects playing the recordings more times will have better recall scores
- *Hypothesis* b_2: within each recording, segments played more times will be better recalled by the subjects
- *Hypothesis* c_2, *or the "a-b-a" strategy hypothesis:* most subjects will follow a three-stage playback strategy: starting by playing the recording once straight through, then working through it in sections, and ending by playing it one last time straight through.

It was also proposed that the data should be analysed for two additional purposes

- To define a suitable taxonomy and terminology for the study of playback strategies
- To identify recurrent playback patterns and compare them to those hypothesised in Experiment 1 for further investigation in Experiment 3.

One particular pattern was found in Experiment 1, and will be tested as a new hypothesis in this experiment:

- *Hypothesis* d_2: subjects make more pauses in the initial segments of recordings than in the final segments.

8.2 Procedure

8.2.1 Data collection:

- *Subjects:*

The subjects in this experiment were 16 undergraduate students at the University of Newcastle-upon-Tyne, taking a beginners' course in Spanish as a foreign language as part of degrees in Sciences, Engineering and other non-language areas in 1995-96. All of them had received between 40 and 60 hours of formal instruction in the foreign language.

- *Materials*:

The two recordings used in this experiment (recordings E and F[22]) were fully scripted. Recording E was an answerphone message, read out relatively slowly by a native speaker recorded in a studio. Recording F was a news report about a football match recorded from Spanish radio and spoken comparatively faster. Both recordings were of a clearly transactional nature as they aimed to cover a given number of specific information items within a given time limit. Therefore they presented little redundancy and it was easy to pinpoint the segments of the recordings in which a particular item was mentioned. Each text was accompanied by a list of comprehension questions (Appendices 3.1 and 3.2). The questions covered all the items of information that were explicitly presented in the recordings. In the case of recording E, these questions were in the same sequence in which these items were mentioned by speakers in the recording, whereas for recording F the questions were less directly related to the sequential structure of the information presented (see p.172 below).

- *Procedure:*

The experiment was conducted in a language laboratory. Both recordings were copied onto the subjects' workstations, after which they were instructed to operate the tape freely for up to 30 minutes while they answered the questions given. They were allowed to tackle the recordings in any order they wished, and also to move from one recording to the other as often as they wanted. Using the laboratory's pairing facility, each work station had been connected to a second machine which recorded all the

[22] See transcripts and translations in Appendices 2.8 and 2.9 respectively.

operations (moves) performed by the subjects on the tape, such as rewinding, pausing, playing back in sections, and so on. The subjects were informed of this and agreed to the procedure. As the answers to the questions were to be used as part of the formal assessment for the course, these test conditions could be regarded as "typical" of the conditions in which listening tests are normally administered for assessment purposes in the institution concerned. After the test, the tapes on which individual moves had been recorded were collected along with the written answers to the comprehension questions.

8.2.2 Data processing:

- *Recall scores*

Written test answers were scored using a checklist of idea-units based on the information that was explicitly given in the recordings. This comprised 20 ideas for recording E (Appendix 4.7), and 15 ideas for recording F (Appendix 4.8).

- *Move logs*

Individual moves made by each subject were transcribed in the form of move logs (see p.125). The methods used for analysing move log data will be described throughout this chapter.

- *Playback frequency:*

The overall correlation between the number of moves made and individual recall scores was calculated. In addition, a comparative measure of playback frequency between successfully and unsuccessfully recalled segments was calculated as follows:

For each idea-unit, the exact segment of the recording in which the relevant information had been presented was identified. In that way, it became possible to count how many times the segment(s) related to a particular idea-unit had been played back by each subject. Two different scores were calculated in this way:

1. *"Recalled" playback score:* Only the ideas that had been *successfully* recalled were counted in this calculation. The number of times that the segments related to each of these ideas had been played back was counted. The values obtained for all "recalled" ideas were then averaged for each subject. This average number of times was termed the subject's "recalled" playback score.

2. *"Not recalled" playback score:* The same calculation was performed again, this time including only the ideas that had *not* been successfully recalled. The resulting average was termed the subject's "not recalled" playback score.

The point of this comparison was to see whether the number of times a particular segment had been played was related to successful recall of the ideas "contained"[23] in the segment in question.

8.3 Results

In order to test the hypotheses formulated at the beginning of the chapter, this section will consider the following aspects of individual playback strategy.

1. *Task sequencing:* the order in which the two recordings were played back (recording E = *"task one"* in the test; recording F = *"task two"*).
2. *Use of straight runs at start and end:* the use of "straight runs" (i.e. the whole recording played back from beginning to end without interruption) as the very first or very last move made while performing the task.
3. *Playback frequency:* the overall correlation between the number of moves made and recall scores, and a comparison between "recalled" and "not recalled" playback scores.
4. *Types of playback patterns:* visual observation of students' move logs, in order to see whether any typical playback patterns could be identified
5. *Position of move clusters in the move logs for recordings E and F.* Move clusters will be defined in section 8.3.4.1 of this chapter (p.167).

8.3.1 Task sequencing

Table 8-1 shows the order in which subjects tackled tasks one and two (recordings E and F respectively, shown in low caps for readability).

Eight subjects (50%) tackled the tasks in sequential order (*"ef"*), exactly as they were presented to them.

[23] See p.118 for a critical discussion of the notion of segments "containing" ideas.

Four subjects (25%) tackled them in sequential order, but then returned to task one and repeated the sequence again (*"efef"*). Two additional cases might also be included in this group (T10 and T05), although they are slightly different:

T10 began by playing both recordings straight through, one after the other without interruption (*"e+f... "*); then returned to task one and proceeded sequentially (*"...ef"*).

Student codename	*Order of tasks*	*Recall score (%)*
T01	efef	66.3
T02	ef	53.1
T03	efef	58.8
T04	ef	80.6
T05	ef(e)f	73.8
T06	ef	82.5
T07	efef	68.8
T08	ef	50.6
T09	efefe	31.9
T10	e+fef	68.8
T11	efefef/	68.1
T12	ef	73.8
T13	ef	80.0
T14	ef	61.9
T15	efef	71.9
T16	ef	68.8

Table 8-1: Task sequencing

T05 appeared to follow an *"efef"* sequence. Yet it is possible that the return to task one (shown in brackets on Table 8-1) could be due to having rewound the tape too far back. When this (accidentally?) happened, the subject might have just decided to allow the tape to play on until it reached the second recording again. This interpretation is based on the fact that only the end of the first recording was played, in one single move, and that no pause was made before proceeding to the second recording.

Two subjects (12.5%) repeatedly switched from one task to the other: T09 (*"efefe"*) and T11 (*"efefef"*). The data collection tape for T11 ends suddenly (shown by a slash after the last " f " on the table), making it difficult to establish what the complete playback strategy would have been in this case.

- *Effect on general recall rates*

The recall scores of the eight subjects who tackled the recordings in the same sequence in which they were presented to them (*"ef"*) were compared with those of the eight subjects who switched from one task to the other (*"efef+"*). No significant difference was found between the two means (t-Test: Paired Two-Sample for means: t=1.07; t critical two-tail=2.36 with p=0.32).

The results are shown on Table 8-2 below.

Task sequence used:		*"ef"*	*"efef+"*
Average recall score (%):	*Mean*	68.9	63.5
	S.D.	12.5	13.5

Table 8-2: Effects of task sequence on general recall

8.3.2 Use of straight runs at start and end

Individual move records were examined in order to establish which subjects had started tackling the tasks by playing the recordings straight through at least once, and which subjects had ended the tasks in this way. The following conventions were adopted:

- Full runs of the recording containing only one pause were counted as straight runs, so long as the tape had not been rewound at any point until the end of the recording (full runs containing more than one pause were not counted as straight runs).
- If the end of the recording was truncated by no more than two segments, the move was still counted as a straight run.
- If the beginning of the recording was truncated no more than five segments (i.e. before the first idea-unit relevant to the test questions was mentioned), the move was still counted as a straight run.

The results of this analysis are summarised in Table 8-3 below (see Appendix 6.1 for the full results).

In 24 cases out of 32 (75%), the task was started with a straight run. This occured in 12 cases with each of the recordings, although the strategy was not used by the same 12 individuals in both tasks.

As for straight runs at the end of tasks, there was a clear difference between recording E (only 5 cases: 37.5% of task one) and recording F (10 cases: 62.5% of task two). Overall, this strategy was used 15 times out of 32 (47% of all cases).

The final straight run was therefore omitted 53% of the time, but even so, a straight run was generally used at the start. There were 10 such cases in which a straight run occurred at the start, but not at the end of the task.

In the second most frequent pattern, no straight run was used, either to start or to end the task. Four such cases (25%) were observed in recording E and 3 in recording F. (18.7%). Over the two recordings, this strategy was used 7 times out of 32 (22%).

The least frequently used strategy (only one case was observed, in recording F) was to end the task with a straight run while having started it with shorter moves. This was not a typical behaviour. Normally if the straight run was omitted at the start, it was also omitted at the end of the task.

Use of straight runs:	*Recording E*	*Recording F*	*Total*
• At start and end of task	5	9	***14***
• Only at start of task	7	3	***10***
• Neither at start nor at end of task	4	3	***7***
• Only at end of task	0	1	***1***

Table 8-3: Use of a straight run as the first / last move of the task in recordings E and F (summary of results).

As for strategy patterns across recordings E and F, the most frequent was to use a straight run at the start of recording E, but not at the end; and then to use a straight run both at the beginning and the end of recording F. Five subjects (31% of the sample) followed this strategic pattern.

Each of the following patterns was only adopted by two subjects (12.5% of the sample in each case):

- Straight run used in recording E: at start and end; in recording F: at start, but not at end.
- Straight run used at start and end in both recordings.
- Straight run used in recording E: at start, but not at end; in recording F: at start and end.

Each of five remaining patterns was followed by only one subject in the sample (a different subject in each case).

In summary:

The use of straight runs as first or last moves of a task in this experiment can be summarised in three general tendencies:

- The most common strategy was to begin a task with a straight run, whether or not the same strategy was adopted for terminating the task.
- The general tendency for task one (recording E) was to use a straight run only at the start, whereas for task two (recording F) it was used both at the start and at the end.
- If a straight run was not used at the start of a task, is was not used at the end of the task either.

8.3.3 Playback frequency

Recall scores did not correlate with the total number of moves made by each subject. The overall correlation between segment recall and playback frequency was also negligeable and failed to reach significance. However, a more detailed analysis of recall in relation to playback frequency did produce significant results.

All the segments played back by each subject were divided into two categories:

1. Segments related to ideas that the subject had successfully recalled (for convenience, these will be referred to as "Recalled" segments)
2. Segments related to ideas that the subject had failed to recall (referred to as "Not recalled")

For each subject, the average playback frequency (the number of times segments had been played back) was calculated separately for *recalled* and *not recalled* segments (see p.158 for details). The results obtained are summarised below (full data in Appendix 6.2):

	Recalled	*Not recalled*
Mean	**9.19**	**8.03**
S.D.	**2.31**	**2.11**

Table 8-4: Average playback frequency of segments containing "recalled" and "not recalled" ideas (recordings E + F)

Except in two cases (subjects T09 and T12, circled on Table 8-5 below), the average playback frequency within each subject's scores is always greater for "recalled" than for "not recalled" segments. The effect is not very large ("recalled" segments were played back only 1.16 times more on average), but a paired t-test shows that this difference is highly significant, with p=0.0005 (t=4.36> t-critical two-tail=2.95).

This confirms the visually clear results illustrated in Table 8-5:

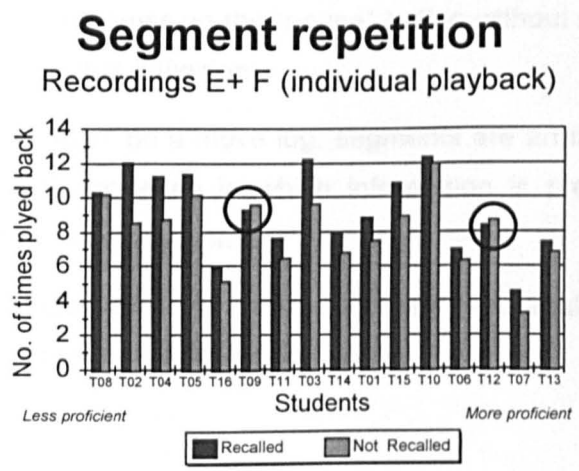


Table 8-5: Segment repetition - Recordings E+F

8.3.4 Types of playback patterns

Much of the analysis of move logs entailed producing a suitable taxonomy and a practical terminology for the description of playback-related strategies. There was also a need for a clear, yet concise notation system to describe the observed strategies. Therefore the first part of this section will take the form of a glossary in which the new categories are defined and the notation system is explained. This system will be used throughout the discussion from this point onward.

8.3.4.1 Local playback patterns: moves, segments and runs

Move:

- *Definition:* see p.125.
- *Visual representation:* see p.125.
- *Notation:* the abbreviation for "move" is a capital M ("3M. "means "three moves"; "M3" means "move number 3").

Moves→	1	2	3	4	5	6	7	8	9	10
Segment 1										
Segment 2										
Segment 3										
Segment 4										
Segment 5										
Segment 6										
Segment 7										
Segment 8										
Segment 9										
Segment 10										
Segment 11										
Segment 12										
Segment 13										
Segment 14										
Segment 15										

Table 8-6: An example of move log

Segment:

- *Definition:*

  Every recording on this study was divided into segments (the shortest units that could be isolated by pressing the "pause" button without cutting words - see p.121 for a discussion of this definition).

- *Visual representation*: on a move log, segments are arranged vertically and follow the chronological sequence in which information is presented in the recording (progressing from top to bottom).

- *Notation*: the abbreviation for "(tape) segment" is a small "s" ("8ss." means "eight segments"; "s8" means "segment number 8").

Run:

- *Definition:* each time the whole recording is played from beginning to end (regardless of any interruptions or repetitions that may occur) is counted as a run. A run normally begins with the first segment of the recording (s1) and ends with the last. However, if other parts of the recording are played after the last segment has been heard, the new run will not begin until s1 is played again. A run can be described as *fast* or *slow*, depending on the number of moves that it took to complete the run. This last feature will be referred to as the *pace* of progression in a given run

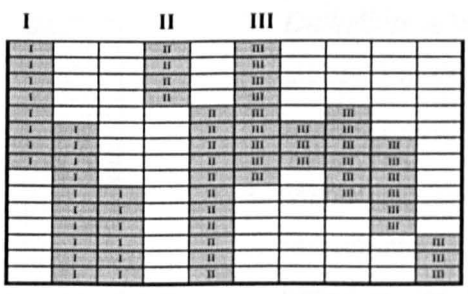


Table 8-7: Three runs

(e.g. run II in the example shown above is faster than run III). The description of a run as "fast" or "slow" is relative to the other runs to which it is being compared. Identical patterns may be regarded either as fast or slow, depending on the pace of the other runs found within the same move log or across a sample of move logs from different subjects.

- *Visual representation*: on a move log, runs can have different shapes depending on how fast the subject has progressed through the recording. A straight run (see below) has the shape of a vertical line and represents the fastest progression possible. As the progression gets slower (i.e. more sections are played repeatedly before moving on) the general line will be broken into shorter moves that are roughly arranged in a diagonal line. The slower the progression, the more this line stretches horizontally. In a few cases the overall pattern of a slow run may not be a diagonal line (if students do not proceed sequentially through the recording), but this is a rare occurrence.
- *Notation*: runs are notated in roman numbers ("II" means "run number two).

Straight run:

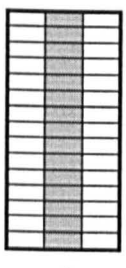


Table 8-8: A straight run

- *Definition:* a run that is made up of one single move. This means that the whole recording has been played back straight through from beginning to end without interruption.
- *Visual representation*: on a move log, a straight move is represented by a vertical line that begins at the first segment and ends at the last. In certain cases, a move in which the first or the last few segments of the recording have been omitted may be still counted as a straight run (see p.161).

Run with breaks (broken run):

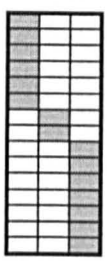


Table 8-9: A run with two breaks

- *Definition:* a broken run is a run in which the tape has been stopped at least once, but has never been rewound to play back previous segments.
- *Visual representation*: on a move log, a broken run is represented by a broken line in which there is no overlap between one move and the next. The points at which the

line is broken represent the pauses that were made.

Run with (single) recaps:

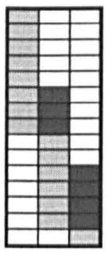


Table 8-10:
A run with two recaps

- *Definition:* a run in which the subject pauses the tape, rewinds it a little bit in order to play back the last few segments again, and then continues onto the next section. The section that is played again may comprise just a few segments (*short* recap), or a significant proportion of the previous move (*long* recap).
- *Visual representation*: on a move log, the parts that are repeated are shown as overlapping lines (shown as shaded areas on Table 8-10.

Run with clusters (clustering recaps):

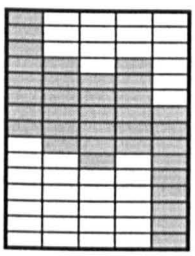


Table 8-11:
A cluster

- *Definition:* in a run with clusters, the same section is repeated more than once before moving on to the next section. The more times a section is played, the slower the progression, the thicker the cluster.
- *Visual representation*: on a move log, this playback strategy produces "clusters" of overlaping lines covering the segments that have been repeated. Cluster boundaries are often difficult to define precisely, and clusters themselves can be produced with varying degrees of overlap between one cluster and the next.

This terminology can be used in order to produce a concise description of the playback pattern applied by a subject to a particular recording. For instance, the first move log shown in this section (Table 8-6, p.165) could be summarised as:

> I=3M (2 overlaps); II=2M (1 break); III=5M (1 cluster around s.6-9).

In terms of pace, run II could be described as the fastest, and run III the slowest, because it takes up five moves, as opposed to two moves for run II.

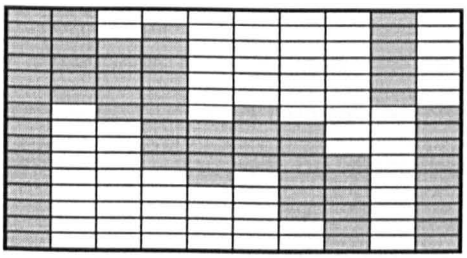
8.3.4.2 Global playback patterns

Visual analysis of the move logs as a whole (i.e. all the moves made by one particular individual while working on a given recording) also showed a number of global patterns that will now be described. Each typical pattern represents a patricular playback strategy.

The a-b-a playback strategy

This pattern was the original hypothesis for this study: subjects begin with a straight run, then progress in clusters through the second run, and complete the task with a final straight run (I=straight; II=with clusters; III=straight). Only one move log (T16, task 1) displayed this pattern as described.

a | ← ← ← b → → →| a →



An example of the a-b-a pattern

In fact, many patterns did approach the description to some degree. A looser description of the a-b-a pattern as: I=fast; II=slower; III=fast would better reflect the actual playback strategies that were observed. The "fast" runs could either take the form of a straight run, or be broken in one or two places. The "slower" run would include either single recaps or clusters of varying paces.

Progression in cycles

One typical variation of the a-b-a strategy may comprise two or more slow runs (instead of only one) between the two fast runs that begin and end the task. In other variations, it may be one of the fast runs that is repeated.

I | II | III

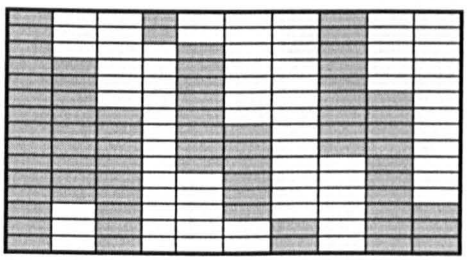


Table 8-12: Progression in cycles

This type of strategy is refered to as a "progression in cycles", and is characterised by a succession of runs of similar type (either fast or slow) at some point in the move log. Table 8-12 shows three slow cycles as they might occur in the middle of a move log.

Fast pace strategy

A significant number of subjects adopted a fast pace strategy, in which progression is mostly in the form of straight moves. Visually, the resulting pattern looks like a compact succession of straight columns on the move log, with only sporadic breaks (if any). Because the moves are longer than in other types of logs, fewer moves are made. Therefore fast pace logs tend to be shorter than others on the horizontal axis.

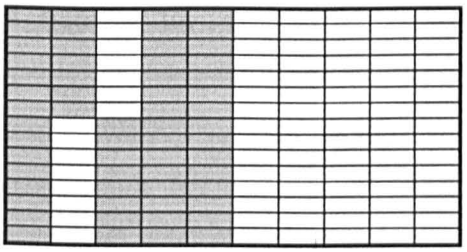


Table 8-13: Fast pace strategy

Slow pace strategy

The slow pace strategy was not very common for the two recordings used in this experiment, but it is worth describing here as it was frequently observed in Experiment 3, and will be discussed in later chapters. Basically, it comprises a succession of short moves, with very frequent recaps. This produces numerous clusters of varying paces and lengths. It reflects intensive work on short sections of the recording at a time, and has the appearance of a series of short moves arranged more or less clearly in a diagonal line as shown on the example.

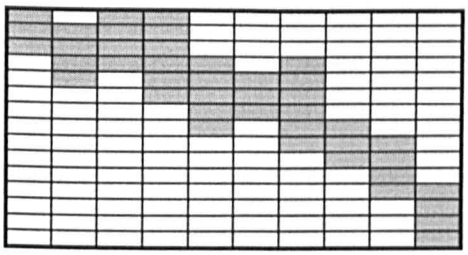


Table 8-14: Slow pace strategy

Top-heavy strategy

As a general tendency, subjects often listen more times to the beginning of the recording than to the end. This tendency is particularly striking in a number of move logs in which there is a much greater number of moves centered around the first few segments of the recording. In these cases, it would seem that a slow pace strategy has been adopted for the beginning of the recording, whereas progression through the rest of the recording is much faster. Visually, the move logs reflecting this strategy have a "top-heavy" appearance, with thick

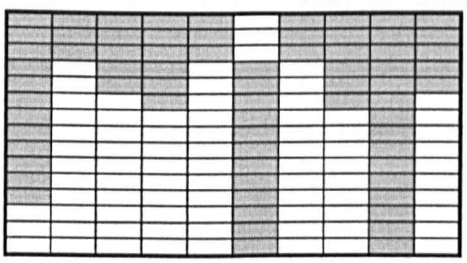


Table 8-15: Top-heavy strategy

clusters (slow progression) at the top and much thinner columns (fast progression) at the bottom.

8.3.5 Position of move clusters in the move logs for recordings E and F

A study of the location of clusters[24] was conducted in order to establish in what sequence the questions had been tackled in recordings E and F.

Unfortunately, precise cluster boundaries are extremely difficult to define. For instance, a fair number of short moves would appear to be the result of tape operation difficulties (subjects trying to find the right spot on the tape)[25]. Spoken utterances are much more difficult to pin down than written sentences or paragraphs, because speech is time-bound and invisible. The lack of interaction with the speaker and the lack of visual channel in audio recordings also causes them to be generally perceived as quite fast, especially in L2 (cf. section 3.2.1). Consequently, human learners cannot be expected to produce neat textbook examples of typical playback strategies. For this reason, the study of move logs focused on global features rather than detail, so that the "noise" resulting from this natural margin of error was minimal and general patterns could begin to emerge. Obviously this approach is valid for hypothesis-generating rather than hypothesis-testing. The suggested patterns were further tested in Experiment 3.

Described below are "first-glance" findings from visual examination of move cluster position in the move logs from recordings E and F. Code numbers beginning with the letter T refer to individual subjects taking part in Experiment 2.

8.3.5.1 Recording E ("task one")

Features of recording E:

1. This recording was spoken at a fairly slow speech rate, a feature that made it easy for subjects to stop between sentences.
2. The comprehension questions in the student's brief followed exactly the same sequence as the information in the recording.

---

[24] The term "cluster" was defined on p. 167.

[25] One could argue that digital audio might provide more accurate data. Yet researchers would then be faced with the dilemma of whether of not recordings should be pre-segmented for the subjects for easy tracking, and if so on what basis.

Playback strategies observed:

1. Recording E was generally tackled in sequence (e.g. subjects T06, T12, T16).
2. One easy section could be picked up while listening, with no visible evidence of focus in the form of a move cluster (e.g. segments related to question 4, passed through by T09, T15 and T03).
3. Often there was a *progressive sequence* through the sections: play sections A+B+C, then B+C+D, then D+E+F and so on. Subjects T04, T13 produced very neat examples of this strategy; and to a lesser extent, T08. Sometimes an intervening break was inserted in order to work on the second task for a while (subject T03).
4. Less charateristic sequences were also found in individual cases (e.g. subject T01).

8.3.5.2 Recording F ("task two")

Features of recording F:

1. Recording F was spoken much faster than recording E, so it was not as easy to stop between sentences.
2. Also, the questions asked in the task brief were not exactly in the same order as the information presented in the text, as shown in Table 8-16.

Playback strategies observed

To avoid repetition, some of the results below will be reported and discussed within the same section.

- *Question 1 ("who are the two teams playing?"):*

No dedicated attention to the segments related to this question was observed (they were always played along with other segments). This is probably because it was very easy to understand and the answer could be picked out straight away (the two idea-units "Madrid" and "Barcelona" were the only two that had a facility index of 1 in this test[26]). This hypothesis is in line with the "facility threshold" hypothesis of Experiment 1 (i.e. no pauses made if the input was too easy or too difficult). It will be tested further in the next experiment (free playback in pairs), in which students' spoken

[26] A facility index of 1 indicates that every subject in the sample scored full marks on these idea-units.

Information in the recording *(shown in their original sequence)*	***Questions in the task brief*** *(see question numbering for sequence)*
1. Time of broadcast	• This information could be mistaken for question 2: "when will the match take place?"
2. Price paid for tickets	• Question 4: "At the beginning of the clip, a figure is given. What is this figure and what does it refer to?"
3. Teams playing	• Question 1: "Who are the two teams playing?"
4. Mood in the city	• Question 5: "What is the general mood in the host city about this match?"
5. Time of match and weather forecast (too close to each other - interwoven - to be played back separately)	• Question 2 (time): "When is the match taking place" • Question 3 (weather): "What is the weather forecast for the match?"
6. Spectacular performance expected.	• Could be associated with question 5 (mood): *"What is the general mood in the host city about this match?"*
7. Teams playing (2nd time mentioned)	• Question 1: *"Who are the two teams playing?"*

Table 8-16: Sequence of information and related questions in recording F

responses when listening to the same segments will either confirm or disprove this interpretation.

- *Questions 2 ("time of match") and 3 (Weather"):*

A few individuals (50%) focused on the segments related to questions 2 and/or 3 in their second cluster (T02, T04, T06, T09, T12, T13). Unfortunately, the answers to both questions could be found in the same segment, at a very high speech rate, which makes it difficult to determine whether the subjects' focus when playing these segments was on "time of the match", "weather" or both.

- *Overall pattern for cluster location:*

The most typical distribution of move clusters through recording F follows a sequence that would appear to be determined by questions 2, 3, and 4, rather than by the original sequence of information in the recording:

1. Cluster 1: located around the opening segments of the recording, which stated the time of broadcast ("It's now 3 o'clock"). It is possible that subjects heard mention of

a time and assumed this ought to be the actual time of the match (which was less salient information and was given later on in the recording). This would mean that the first cluster was related to question 2 ("When is the match taking place?"). Spoken responses from Experiment 3 (free playback in pairs) should confirm whether this was indeed the case.

2. Cluster 2: located around the segments related to the time of the match and weather forecast (the focus of questions 2 and 3 respectively). Given that in the previous cluster (located at the beginning of the recording), most subjects seemed to have been already looking (mistakenly) for the answer to question 2, it is necessary to establish whether, in this second cluster, those subjects were actually trying to answer question 2 or question 3. This will only be possible in the next experiment.
3. Cluster 3: located around the segments related to the price paid for the tickets (associated with question 4).

No cluster was observed in relation to question 5 ("mood in the city"). The saliency of the key word "pasión" could account for this, because it could be recognised without the need for repetition. The same could apply to question 1. The lack of a specific cluster suggests that the question was answered straight away and did not require special attention. This interpretation will be examined again in Experiment 3.

- *Progressive sequence:*

As well as the overall pattern described above, a progressive sequence similar to the one observed in recording E was also observed here (subject T16, who also used this technique in recording E).

In summary:

Assuming that questions 1 and 5 were answered "on the fly" (because they were the easiest), the playback patterns observed here seem to reflect a sequence corresponding to the sequence in which the questions were presented (rather than the original sequence of the input). This was to be checked against protocol data from Experiment 3. However, a few subjects did follow a sequential pattern of progression through this recording.

8.4 Discussion

This section will discuss the results obtained in Experiment 2, with particular attention to task sequencing, playback frequency and playback patterns. The order in which clusters were found in the move logs will be discussed as an aspect of task sequencing. The use of straight runs will be examined under the section devoted to playback patterns.

The provisional conclusions emerging from this discussion constitute a series of hypotheses to be tested in Experiment 3.

8.4.1 Task sequencing

All students, without exception, tackled the two recordings in the same order in which these were presented to them. The fact that only one out of the 16 subjects listened to both recordings straight through before starting to work on the first one indicates that this behaviour was not related to the relative difficulty of the two recordings, but rather to the sequence in which they were presented. Even if the second task had been significantly easier than the first, only the student who began by playing both recordings would have known this in time to make a strategic decision regarding task sequencing. In a test situation, such a lack of planning could have adverse consequences on time management (for example if a student was to discover half way through the test that the second task was considerably more demanding than the first, and that more than half the time allowed had already been spent on the first recording).

The strategy consisting of using a straight run through the entire input before commencing a listening task is only possible with audio-recorded input. It can be interpreted as a metacognitive PL strategy related to planning. It also appears to be largely underused. Strategy-training issues will be discussed in chapter 10.

Recall rates of students who had tackled the tasks in strict sequence (i.e. first E, then F) were not significantly different from those of students who returned to the first recording again after having worked on the second recording for a while.

As for the order of clusters within each of the recordings, the differences observed suggest that subjects tend to proceed in the same sequence in which information is presented in the input (as they appeared to do with recording E), *unless* their task brief includes questions that do not follow the same sequence as the recording. In

that case they follow the sequence of the questions rather than the original sequence of information in the recording (as they appeared to do with recording F). However, this hypothesis was based on a coarse visual examination of cluster location in the move logs. The evidence was problematic, given the difficulty in establishing clear boundaries for move clusters or precise relations between certain segments and their related questions. The latter could only be achieved by means of protocol analysis (see Experiment 3).

It must be pointed out that cassette is a sequential medium by its very nature. The use of a medium allowing direct access to particular sections might, in theory, generate fewer sequential patterns. However, naturally occurring speech is itself produced and received sequentially. This feature, rather then the actual medium used for PL, is most likely to account for the regular occurrence of sequential PL strategies.

8.4.2 Playback frequency

Segments that were best recalled were also those that were played most often. There are four possible explanations to this phenomenon:

(a) Recall is better because greater exposure to input provides more opportunity for bottom-up processing. If this is the case, Experiment 3 should show that the gain observed from one time to the next is mostly related to decoding strategies.

(b) Recall is better because increased processing time provides more opportunity for top-down processing. If this is the case, Experiment 3 should show that the gain in recall observed from one time to the next is mostly related to inferencing strategies.

(c) Recall is better because the subject is choosing to focus on segments that s/he can understand better in the first place. If this is the case, Experiment 3 should show that a number of recognised items related to the most often repeated segments are present from the very first hearing of these segments.

(d) More than one of the causes listed in these three hypotheses could operate together, in a particular combination determined by the nature of the input, the task, and subject proficiency.

8.4.3 Playback patterns

When categorising specific playback patterns, the underlying assumption is that each of these patterns relates to a particular playback strategy. Such playback strategies are interpreted as task-specific instances of metacognitive strategies that subjects use in order to organise the cognitive processes taking place during the task.

This section will examine hypothetically what metacognitive strategies could be related to the playback patterns that were identified in Experiment 2. It will also outline the type of results that would either validate or invalidate these hypotheses in Experiment 3.

8.4.3.1 Progression in cycles

When several runs of similar type (either fast or slow) occur in succession, this is described as a progression in cycles. One possible explanation for this pattern could be that the subject has not fully completed one stage at the end of the first run of this type. The working hypothesis here would be that subjects continue using the same type of playback strategy until they are ready to proceed to the next stage. This may happen either when they have completed their goal for the current stage (for instance, they stop listening in straight runs and begin to work in slow runs once they have figured out how many sections there are in the recording), or when they consider that they have exhausted the information obtainable by the current means.

If this hypothesis is valid, the protocol transcripts of Experiment 3 should show that when the pace of progression changes, the type of processes recorded also changes.

An opposite hypothesis could also be formulated, whereby stages do overlap each other, even within a single cycle of fast or slow runs. In such case one would observe different types of processing taking place between contiguous runs of similar pace.

8.4.3.2 The a-b-a strategy hypothesis

The fact that a few "fast-slow-fast" patterns approaching the predicted a-b-a pattern were actually found in Experiment 2 supports the hypothesis that the a-b-a pattern might indeed be related to an existing strategy. This will be referred therafter as the *"a-b-a" strategy hypothesis*, or the *"a-b-a" hypothesis.*

Assuming that the previous hypothesis (whereby change of pace means change of processing strategy) is true, three hypothetical strategies could be associated with each of the three stages of the a-b-a pattern.

1. *Fast progression at the beginning of a task:*

   Hypothesis: In cases where a fast cycle at the start of a task is used, one or more of the following strategies will be observed:

   - getting acquainted with the speaker's accent and delivery
   - identifying the structure of the recording,
   - activating relevant schemata,
   - locating information that "sounds" important.

2. *Slow progression*

   Hypothesis: When slow progression is used, the strategies observed will be mostly related to decoding and local inferencing, involving controlled attention to specific sections of the input.

3. *Fast progression at the end of a cycle*

   Hypothesis: When a fast progression occurs at the end of a cycle, the strategies observed will be predominantly related to hypothesis-checking.

8.4.3.3 Use of straight (or fast) runs in the middle of the task

Subjects who proceed through the entire task in fast runs may be doing so for two reasons.

The first possible reason is that they are allowing the tape to play on, simply because the recording is too difficult for them and they do not believe that interruptions would make the processing any easier. This would be consistent with some of the remarks made by subjects in Experiment 1, suggesting that is was pointless to use pauses in difficult recordings. In such cases, the lack of breaks or recaps would indicate that no processing is taking place. In Experiment 3 protocols, this ought to be reflected in an absence of factual information reported during or after the fast runs in question.

The second reason for proceeding in fast runs even after the initial stages of a task is quite the opposite: if the decoding process is sufficiently automatised, it can take place on-line (i.e. while the tape is playing), and therefore there is no need to use any breaks or recaps. Again, this hypothesis is consistent with some of the comments

made by students in Experiment 1, who argued that there was no need to stop the tape if the recording was easy. Indeed the results showed that none of the subjects in group C inserted more than one single pause in recording A2 (the easiest of the three recordings that were used).

The two hypotheses to be tested in relation to fast overall progression could therefore be formulated as follows:

1. In difficult passages where a fast overall progression occurs, little processing will be found to take place.
2. In easy passages where a fast overall progression occurs, most decoding processes will take place automatically.

These hypotheses were to be tested in Experiment 3, using recordings of varying difficulty as input.

8.4.3.4 Progression in slow runs

When a subject chooses to play back short segments repeatedly, this could indicate that controlled processing of short units of input is taking place. Experiment 3 was expected to confirm this, as explained in section 8.4.3.2 above.

8.4.3.5 Top-heavy patterns

There are two possible reasons that could explain why subjects would pay more attention to the beginning of a recording than to the end:

When a particular passage is selected and edited to constitute a "recording" (be it by a language teacher, a researcher or a media programme editor), it is generally selected so that it is a self-contained unit. We also know that, when speakers introduce "new" information, they take into account what information has already been "given" to the hearer (Brown and Yule 1983a; Halliday and Hasan 1985). In order to understand a message, listeners must first identify the key parameters that will be used in subsequent referential expressions (e.g. "the girl" assumes that the hearer already knows what girl we are talking about) and anaphoric reference (e.g. *"she"*=the female we introduced earlier) (Brown and Yule 1983a:208). This would imply that accurate decoding of the initial segments of a recording is an essential first step in order to understand the rest of the message.

McCarthy (1991) examines a number of formal schemata underlying the structure of typical text types. His account, which is largely based on written examples, comprises discourse patterns such as "problem-solution"; "general-specific", "claim-counterclaim", "question-possible-answers-chosen answer" (see also p.44). Similar patterns also occur in the spoken language. In a study of the comprehension by Hong Kong undergraduates of a lecture in English based on the formal schema "situation-problem-solution-evaluation", Tauroza and Allison (1994) found that listeners try to recognise patterns that are familiar to them in the incoming message to assist comprehension. Given that it would make little sense to examine the solution to a problem without first knowing what the problem actually is, one would expect that the subjects in the present study should spend more time on the beginning of recordings, in order to establish what the leading idea for the whole recording might be.

Protocol transcripts of Experiment 3 will be searched for any explicit remarks suggesting that this type of strategy is being used.

8.4.4 Conclusions of Experiment 2

In the light of the previous discussion, the original hypotheses for the present study will be reviewed.

1. *Hypothesis a_2:*

   This hypothesis proposed that within the sample, subjects playing the recordings more times would have better recall scores. This could not be confirmed as there was no significant correlation between overall recall and number of moves made.

2. *Hypothesis b_2:*

   This hypothesis proposed that within each recording, segments played more times will be better recalled by the subjects. In this case, the hypothesis was confirmed: a small but highly significant difference was observed between average "recalled" playback scores and average "not recalled" playback scores.

3. *Hypothesis c_2:*

   The "a-b-a" strategy hypothesis claimed that most subjects would follow a three-stage playback strategy, starting with playing the recording once straight through, then working through it in sections, and ending by playing it one last time straight

through. The results were not conclusive, although a number of "fast-slow-fast" patterns approaching the hypothesis were found.

4. *Hypothesis d_2:*

   This hypothesis claimed that subjects would make more pauses in the initial segments of recordings than the final segments. The hypothesis was confirmed in a number of cases, although not across the entire sample.

5. *Data analysis to define a suitable taxonomy and terminology for the study of playback strategies*

   Observation of PL strategies was operationalised by defining a series of "local patterns" that are described in section 8.3.4.1 of this chapter. They are based on formal examination of the runs produced in the move logs, and may consist of straight runs, runs with breaks, recaps and move clusters.

6. *Data analysis to identify recurrent playback patterns and compare them to those hypothesised in Experiment 1 for further investigation in Experiment 3.*

   These patterns were described under the term "global patterns" in section 8.3.4.2 of this chapter, and discussed in the previous section. They are based on formal examination of the pace of progression throughout each single move log. Such patterns include a-b-a (or fast-slow-fast) patterns, progressions in cycles, fast progressions, slow progressions, and top-heavy patterns.

The next chapter discusses the final experiment of this study. The aim of Experiment 3 was to test the hypotheses generated in the two experiments reported so far and to explain some of the patterns observed.

9. Experiment 3 - Free playback in pairs

9.1 Hypotheses to be tested

The hypotheses that Experiment 3 was designed to test had been initially proposed in the form of two very broad premises (see section 5.3):

A. Similarities in playback patterns will be observed across populations performing identical tasks, and such trends will be different for different types of input.

B. Specific relationships can be established between a given set of playback patterns and specific cognitive, metacognitive and social-affective strategies.

A third hypothesis was also suggested, assuming that A and B were true: "Given identical tasks, individual variables (e.g. proficiency, learning style, motivation) will affect the playback strategies that subjects use." However, it was not envisaged to include this qualitative analysis of individual differences in the present study.

On the basis of results from experiments 1 and 2, hypotheses A and B were operationalised by means of a series of specific hypotheses that are listed below.

- **Testing hypothesis A**

In order to test hypothesis A, the move logs produced for each of the recordings were compared, with particular attention to the two following aspects:

1. Pace-related variables (the number of moves and runs made during the task as well as their length and distribution).
2. Repeat rates (the number of times each segment of the recordings had been played by the subjects).

- **Testing hypothesis B**

This hypothesis was tested through a selective analisis of protocol data, focusing on the features that were directly relevant to the following sub-hypotheses:

Pace of progression

1 Fast progression at the beginning of a task

In cases where a fast cycle at the start of a task is used, it will represent one or more of the following strategies:

1.a getting acquainted with the speaker's accent and delivery

1.b activating relevant schemata,

1.c locating information that "sounds" important.

2 Fast progression at the end of a task

2.a When a fast progression occurs at the end of a task, the strategies observed will be predominantly related to hypothesis-checking.

2.b Whatever the pace, the last run in a task will be predominantly related to hypothesis-checking.

3 Fast overall progression

3.a Fast progression overall indicates that little processing is taking place.

3.b Fast progression overall indicates that most decoding processes are taking place automatically.

4 Slow overall progression

During a slow overall progression, controlled bottom-up processing is more intensive than when a faster progression is used.

Global playback patterns

5 The a-b-a (fast-slow-fast) pattern

In cases where an a-b-a pattern is observed, the three stages of the cycle will correspond respectively to the strategies proposed as hypotheses 1, 4 and 2 above.

6 Top-heavy patterns

In cases where a top-heavy pattern is used, one of the following strategies will be observed while subjects play the beginning of the recording:

6.a Identifying a familiar formal schema (e.g. problem-solution) and trying to interpret the content of the first component of such pattern.

6.b Identifying referrents that are used in subsequent referring expressions and anaphoric reference.

Playback sequence

7. When the listening task is accompanied by a set of written questions, subjects will proceed through the task according to the order in which the questions are presented, rather than the order in which information is heard in the recording.

Playback frequency and recall

8 Segments that are best recalled are also those that are played most often (see results of Experiment 2). One of the following hypotheses is expected to account for the phenomenon:

8.a Recall is better because greater exposure to input provides more opportunity for bottom-up processing. If this is true, Experiment 3 will show that the gain observed from one time to the next is mostly related to decoding strategies.

8.b Recall is better because increased processing time provides more opportunity for top-down processing. If this is true, Experiment 3 will show that the gain observed from one time to the next is mostly related to inferencing strategies.

8.c Recall is better because the subject is choosing to focus on segments that s/he can understand better in the first place. If this is true, Experiment 3 will show that a number of recognised items related to the most often repeated segments are present from the very first hearing of these segments.

8.d More than one of the causes listed in these three hypotheses could operate together, in a particular combination determined by the nature of the input, the task, and subject proficiency.

Social-affective strategies

9 Social-affective strategies will overtly affect the playback strategies that subjects use.

9.2 Procedure

9.2.1 Data collection

- *Subjects*

The subjects in this experiment were 35 undergraduate students at the University of Newcastle-upon-Tyne, taking a beginners' course in Spanish as a foreign language as part of degrees in Sciences, Engineering and other non-language areas. All of them had received between 40 and 60 hours of formal instruction in the foreign language. Two successive cohorts took part in the experiment, one in 1995-96 (8 students, all tested in the same session using recording D), and the other in 1996-97 (25 students). For the second cohort, three different sessions were held, using recordings C1, B1 and F respectively. Students attending each session were paired up on the basis of the listening marks obtained at their latest course assessment test, so that subjects working together had similar levels of proficiency. In sessions attended by an odd number of students, the subject with the highest proficiency level was asked to work on his own.

Attendance at the sessions was distributed as follows[27]:

Recording D:	8 students, working in 4 pairs.
Recording C1:	11 students (4 of which had also done recording F), 5 working in pairs and 1 working alone.
Recording B1:	11 students (2 of which had also done recording F), 5 working in pairs and 1 working alone.
Recording F:	11 students (6 of which also attended other sessions as stated above), 5 working in pairs and 1 working alone.

- *Materials*

The recordings used were recordings B1 and C1 (already used in Experiment 1, see p. 138), recording F (already used in Experiment 2, see p.157), and a fourth recording that had not been used before: recording D. A full transcript of the recordings that are summarised below can be found in section 2 of Vol.2. These comprised:

Recording D:	A man describing how, while he was at a party, somebody came to warn him that his house had been burgled, and what he found when he got there (recorded monologue, unscripted, spoken by a teacher - the owner of the house - at moderate/fast speech rate).
Recording C1:	A local TV presenter talking about himself, his job and how the local people perceive him (unscripted interview, fast speech rate).
Recording B1:	An overweight woman asking a doctor for advice about her eating problem (unscripted role-play recorded by two teachers, moderate speech rate)
Recording F:	A sports correspondent reporting on a football match that will take place the day after (authentic radio broadcast, scripted, fast speech rate). This recording was the only one that was presented with a set of written questions.

- *Procedure*

All sessions took place in a language laboratory in which the "pairing" function had been enabled. Each pair of students was allocated two machines. Both students sat in front of one machine (the "player"), of which they were given total control.

[27] For consistency, recordings will be listed always in the same order throughout the present chapter. Although it is not alphabetical, this order corresponds to the order in which move logs and protocol transcripts were sequentially numbered for data processing.

Meanwhile, the second machine was set to "record" mode, so that it captured both the output from the player (all the playback moves made by the subjects) and the students' conversation as they worked through the task together. This second machine was allowed to run uninterruptedly until the task was completed.

Students left without a partner were instructed to talk through what they were doing and say what they could understand, although no training in think-aloud techniques was provided prior to the task itself. Apart from this, the laboratory set-up for individual subjects was the same one as for students working in pairs. Although this procedure was not expected to produce the same type of information produced by students working in pairs, it was decided to collect data from individual students, and take the different conditions into account when analysing the results [28].

Subjects were asked to work through the recording with their partners for 15 minutes. For recordings B1, C1 and D, they were instructed to "try to understand as much as they could together". Note-taking was not allowed during that time, in order to elicit as much spoken output as possible.

In the case of recording F, each pair was given a set of comprehension questions at the beginning of the task (the same questions that were used in Experiment 2 - see p.172) and were simply asked to answer them together, writing down information as and when they wanted.

9.2.2 Data processing

Move logs for the 22 recordings using the same technique that had been used in Experiment 2 (p.125), were produced. These can be found in section 7 of Vol.2).

The 22 recordings made during the sessions (i.e. the pairs' conversations, or think-aloud protocols for those working individually) were fully transcribed (see Vol.2, section 9), producing the following sample:

- Recording D: Protocols #1 to 4 = 4 pairs
- Recording C1: Protocols #4 to 10 = 5 pairs + 1 individual (#10)
- Recording B1: Protocols #11 to 16 = 5 pairs + 1 individual (#16)
- Recording F: Protocols #17 to 22 = 5 pairs + 1 individual (#22)

[28] These cases were discarded from analysis for some of the hypotheses tested. This is explicitly stated in the relevant sections.

The protocol transcripts were then encoded following the conventions described in section 6.2.2.2 (p.128), and analysed in terms of specific behaviours that were associated with particular processes or strategies.

9.3 Results

This section will review the evidence that was found in the move logs and protocol transcripts in relation to each of the two general hypotheses formulated at the beginning of this chapter.

- *Notation conventions*

Hereafter, protocol numbers will be preceded by the symbol #. Quotes and claims related to specific protocol transcripts will be followed by a reference code such as #20/8, in which the first number refers to the move log or protocol transcript from a particular pair (pair no. 20 in this example), and the number after the slash indicates the move in which the information can be found (here, move 8). Full explanations of the abbreviations used in this section are given in sections 8.3.4.1 and 8.3.4.2.

9.3.1 Move log analysis

The first hypothesis to be tested in this study was Hypothesis A, which states that similarities in playback patterns would be observed across populations performing identical tasks, and that such trends would be different for different types of input. In order to test this, move logs for recordings D, C1, B1 and F were examined from two different viewpoints:

1. *Pace-related variables*
   (number and length of moves made, number of runs, etc.).
2. *Repeat-rates*
   (how many times each pair played each of the segments in the recording)

The results from this study of move log data were intended to help identify which of the protocols (or which of their features) represented instances of the specific strategies examined as part of Hypothesis B. In other words, before one can explain qualitatively what a "top-heavy playback strategy" or a "fast overall progression" might indicate in terms of processing strategies, one must know which pairs actually used such playback strategies in the first place.

9.3.1.1 Pace-related variables

Individual scores were calculated from each move log for a number of pace-related variables comprising: the number of segments played, the number of moves made, the number of runs made, the average number of segments in each move, and the average number of moves in each run. Mean scores from each of the four recordings were calculated for these variables. The results are summarised in Table 9-1 below.

The results of an initial ANOVA test showed significant differences in mean scores for 4 of the 5 variables tested, with significance levels of $p<.01$ for "segments", "moves" and "segments per move", and $p<.02$ for "runs". Mean score differences for the variable "moves per run" were found to be just under the .05 significance level, with $p=.054$.

Differences within each variable were further tested using a post-hoc Newman-Keuls test, which pinpointed significant results for the values that are indicated with asterisks on the table. These will now be examined in detail.

- *Number of segments played*

This value gives a global measure of the total amount of input to which the subjects chose to be exposed while performing the task. The only statistically significant difference was found in recording C1, in which a much larger number of segments was played than in any of the other recordings.

- *Number of moves made*

This value indicates how many times the subjects pressed the "stop" button throughout the whole duration of the task. Recordings D and F scored values close to the average (14.4 moves) which were not significantly different from each other. Recordings C1 and B1 scored extreme values (26 and 4 moves respectively), which were significantly different from each other. In other words, subjects stopped very often on recording C1, and very rarely on recording B1.

- *Number of runs made*

This score indicates how many times the subjects listened to the recording from beginning to end, regardless of the amount of stopping and rewinding made in between. The score for recording F was found to be the highest of all, although the difference was only found to be significant in comparison with C1 and B1. This means that the subjects listened to the whole of recording F more times than they did to the

whole of B1 and C1. The difference between C1 and B1 themselves was not statistically significant.

Pace-related variables	*Recording →*	D (N=4)	C1 (N=6)	B1 (N=6)	F (N=6)	OVERALL (N=22)
No. of segments played	Mean	**160**	**303****	**151**	**198**	**207**
("segments")	*Std. Dev.*	*32*	*90*	*59*	*52*	*87*
No. of moves made *("moves")*	Mean	**12.0**	**26.0***	**4.0***	**14.8**	**14.4**
	Std. Dev.	*6.3*	*14.6*	*1.2*	*8.5*	*12.0*
No. of runs made *("runs")*	Mean	**3.5**	**2.7***	**1.8***	**5.0***	**3.2**
	Std. Dev.	*1.0*	*1.2*	*0.7*	*2.5*	*1.9*
Avge. no. of segments in each move	Mean	**15.0**	**14.8**	**40.6****	**15.6**	**22.1**
(s/M: "segments per move")	*Std. Dev.*	*4.2*	*8.2*	*20.3*	*6.7*	*16.2*
Avge. no. of moves in each run	Mean	**3.6**	**12.5**	**2.4**	**4.5**	**5.9**
(M/run: "moves per run")	*Std. Dev.*	*2.3*	*10.5*	*1.0*	*5.4*	*7.1*

Key: * *Mean scores that differ significantly (p<.05) from some of the other scores.*

** *Mean scores that differ significantly (p<.05) from all the other scores.*

*No asterisk: No significant differences found (except those with ** scores).*

Table 9-1: Pace-related variables for recordings D, B1, C1 and F

- *Segments per move*

The average number of segments in one move tells us to how much input subjects listen in one go. The higher the value, the longer the section played and the faster the progression. Recording B1 was significantly above all the others with a value of 40.6 s/M (double the average), meaning that subjects progressed twice as fast through this recording as they did through the others. The remaining three recordings scored very similar values to each other and any differences found between them were not significant.

- *Moves per run*

The average number of moves per run indicates how much stopping it took each time in order to get from the beginning of the recording to the end. Although noticeable differences were observed (recording C1 scored much higher than the others, suggesting there were more interruptions within each run), these differences were not significant enough to be considered: $F(3, 18)=3.0713$; $p=0.054$.

9.3.1.2 Repeat rates

- *Repeat-rate tables*

The *repeat rate* (RR) of a segment is the total number of times this segment has been played back while performing the task. RR values were calculated for each segment in each individual move log. The scores were then grouped together by recording, and are shown on four repeat-rate tables, hereafter called "RR tables" (see Vol.2, Appendices 8.1 to 8.4). Average RR values from all the subjects using the same recording were also calculated for each segment (i.e., for each row on the RR tables).

In order to identify which segments had been played back significantly more than the others, a simple calculation was performed for each of the 22 move logs: first, the mean and standard deviation of RR's over the whole recording were calculated; then, on the basis of these two values, "high" RR segments were defined as those for which the RR was greater than the mean by at least one standard deviation. These segments are shown in shaded boxes for each move log on the RR tables.

The concept of "top-heavy playback pattern" was operationalised in the form of the following definition: a top-down pattern is deemed to occur when all the high RR segments found in the move log are located at the very beginning of the recording. If any segments located in other parts of the recording also have high RR's, the playback pattern is no longer regarded as purely top-heavy. Sporadic occurrences of one or two low-RR segments within a large cluster of high-RR segments are accepted.

According to this definition, the following move logs display top-heavy patterns:

- Recording D: 4 cases out of 4: #1, #2, #3, #4
- Recording C1: 0 cases out of 6.
- Recording B1: 1 case out of 6: #14
- Recording F: 2 cases out of 6: #19, #22

- *Segment number/RR correlation*

In order to establish whether there was a general tendency to decrease the RR gradually throughout the recordings even in cases that did not display "pure" top-heavy patterns, the following analysis was carried out.

The reference numbers originally given to each segment were used in order to examine whether there was a correlation between a segment's RR and its reference number. Since segments are numbered sequentially, low reference numbers belong to segments located at the beginning of the recording, and higher segment numbers are located towards the end. Therefore a positive correlation would indicate that RR increases as the recording progresses (the end of the recording is played more times than the beginning), whereas a negative correlation would indicate the opposite. Pearson correlation coefficients for these variables were calculated and are shown on the RR tables under the label "s#/RR".

The results for the four recordings were all highly significant at $p<.001$:

Recording D	Recording C1	Recording B1	Recording F
r=-0.96	r=-0.93	r=0.53	r=-0.82

As expected, recordings D and C1 produced high negative correlations, indicating that there was a tendency for the RR to decrease through the sequence of those recordings. A similar, though considerably weaker tendency was observed for recording F. The one recording that was clearly disctinct from the others was B1, which was the only one to produce a positive figure. In any case, the very low index of 0.53 (however significant its *p* value) indicates that there is no consistent overall tendency for this recording.

Given that a number of features were found to be specific to particular recordings (as shown on Table 9-2 below), hypothesis A is seen to be confirmed.

9.3.1.3 Summary and discussion

The findings presented in this section have been summarised on Table 9-2 below, which also presents global trends for two additional features to be discussed in the next section: the use of fast progressions at the beginning of the task, and whether or not the task was completed in the time available.

	Recording D *Burglary at the speaker's house*	***Recording C1*** *Local TV presenter interviewed*	***Recording B1*** *Fat woman at the doctor's*	***Recording F*** *Forthcoming football match*
Segments played		More than in any other recording		
Moves made	Average number	Many	Few	Average number
Runs made		About half as many as in F	Less than half as many as in F (incl. 2 out of 6 cases completed in only 1 run)	More than in C1 or B1
Length of moves used			More than twice as long as those made in any other recording	
Overall playback pattern	All top-heavy		All fast progressions	
Repeat-rate throughout the recording	Clearly decreasing	Clearly decreasing		Decreasing slightly
No. of cases in which the task was completed	2 out of 4	1 out of 6 (deliberate termination)	All completed	All completed
Use of a fast run to start	All start fast			All start fast

Table 9-2: Summary of global playback patterns found in recordings D. C1, B1 and F

A number of tendencies emerge from the comparative study of global patterns in the four recordings used. These will be better explained once the protocols have been analysed and specific relations can be established between playback strategies and listening strategies.

At this stage, the preliminary findings can be used in order to identify certain phenomena that need to be accounted for in the next section.

The most extreme differences in global playback patterns were found between recordings C1 and B1: when listening to C1, subjects used a comparatively small

number of runs, but a comparatively large number of moves. In other words: they gave themselves plenty of exposure to the input, but proceeded very slowly through the task. In contrast, the average number of segments per move was significantly longer for recording B1 than for any of the other recordings. This indicates that the pace of progression was significantly faster in B1 than in recordings D, C1 or F.

One explanation for this (which will be discussed in the next section) could be that higher levels of controlled bottom-up processing are required for recording C1, for reasons that should become apparent through protocol analysis.

On the other hand, when listening to B1, subjects used a comparatively small number of runs, and also a comparatively small number of moves. In other words: they gave themselves less exposure to the input and proceeded quite fast through the task. This strategy implies a lack of controlled processing, which could mean either that no processing at all is taking place, or that all processing required is actually taking place automatically. The latter is the preferred explanation in the case of recording B1, given that the recording was relatively easy (familiar topic, slow speech rate, clear schema obvious from the opening utterance), but only protocol analysis will be able to confirm or reject this hypothesis.

Another clear feature revealed in this study is the presence of top-heavy patterns in all move logs produced for recording D. This requires specific attention and will be discussed in the light of protocol analysis results once these have been examined in detail.

(Minimum) ← **Difficulty / RR effect** → *(Maximum)*

Recording B1	***Recording F***	***Recordings D and C1***
Easiest:	*Moderately difficult:*	*Most difficult:*
– Few moves made – Task completed in all cases	– Average number of moves – Task completed in all cases	– Average/Large number of moves – Task <u>not</u> completed in most cases
• **RR effect:** No correlation between playback frequency and segment number	• **RR effect:** Moderate decrease in playback frequency as recording progresses	• **RR effect:** Clear decrease in playback frequency as recording progresses

Table 9-3: Grading of difficulty in the four recordings used

The overall tendency for a gradual decrease in playback frequency - or repeat rate (RR) - throughout the recordings may or may not be due to similar causes. At this stage, all that can be noted is that such a tendency would appear to be proportional to the relative difficulty of the recordings, as illustrated in Table 9-3. The notion of

"difficulty"[29] is interpreted here in terms of a combination between the number of moves made (how many stops and starts the subjects found necessary) and whether or not the task was completed within the time allowed. According to this, the four recordings could be graded from B1 (the easiest) through F (moderately difficult) to D and C1 (the most difficult).

This is consistent with some of the input features present in the recordings, which are summarised in Table 9-4.

	Recording B1 *Fat woman at doctor's*	***Recording F*** *Forthcoming football match*	***Recording C1*** *Interviewing a local TV presenter*	***Recording D*** *Burglary at speaker's house*
Nature of Input: See Brown and Yule's (1983b) categories (p.50).	*Static* (description /instruction)	*Static* (description)	• Part 1: *Static* (description) • Part 2: *Abstract*	*Dynamic* (narration)
Speech rate	Moderate.	Fast	Fast	Moderate/fast
Orality	Unscripted role-play by two language teachers	Scripted report. from off-air radio	Unscripted interview	Unscripted monologue
Cues initially available for schema activation	Opening utterance: *"Doctor I have a problem"* (s.1), followed by *"I eat too much"* (s.3). → At doctor's surgery re. overeating.	Five comprehension questions given prior to listening → Forthcoming football match.	First few turns in the recording (ss.1-12): *"My name is..."*, *"Ilive..."*, *"What's your occupation?"* → Interview, speaker's personal details.	None. The opening segments (ss1-6) only present a few familiar words, none of them relevant to the main "burglary" topic

Table 9-4: Input features in Experiment 3

Again, protocol analysis should help confirm or reject this interpretation. It should also determine whether or not this is related to the same phenomenon causing the consistent occurrence of top-heavy patterns in recording D.

Given the reliance of any further discussion on protocol analysis results, these are examined in the next section. Discussion of global patterns in the four recordings will be resumed at the end of this chapter, in the light of the findings reported below.

[29] This is a preliminary definition drawn up for operational purposes. A more comprehensive analysis of task difficulty will be presented at the end of this chapter (p.253). See also previous discussion in section 2.2.3.

9.3.2 Protocol analysis

9.3.2.1 Fast progression at the beginning of a task

There were 15 instances of fast runs at the beginning of the task:

- Recording D: 4 cases out of 4 Protocols #1, #2, #3, #4.
- Recording C1: 2 cases out of 6 Protocols #6 and #7.
- Recording B1: 3 cases out of 6 Protocols #13, #15 and #16 (but all progressions were fast overall).
- Recording F: 6 cases out of 6 Protocols #17, #18, #19, #20, #21, #22.

Three sub-hypotheses had been formulated in relation to fast progressions at the beginning of a task, whereby listeners using this playback strategy could be doing one or more of the following:

- *Hypothesis 1a:* getting acquainted with the speaker's accent and delivery.
- *Hypothesis 1b:* activating relevant schemata,
- *Hypothesis 1c:* locating information that "sounds" important.

9.3.2.1.1 Reactions to the speaker's accent and delivery

Results

The relevant protocols were examined for each of the four recordings and showed the following results:

- *Recording D (burglary):*

The two pairs who began the task with a straight run displayed an immediate reaction to the input: "I don't really quite catch a lot of it to be honest" (#1/1); "What a relief it stopped!" (#4/1). The two remaining pairs (#2 and #3) did not begin with straight runs, but their first runs were comparatively fast (2 and 3M respectively). They both interrupted their first move with immediate reactions: "Oh dear!" (#2/1); "Stop it. Stop that!" (#3/1), which could even continue in the second move "Huh! Bugger me!"(#2/2).

- *Recording C1 (interviewing a TV presenter):*

Of the two pairs who began the task with fast runs (#6 and #7), only one reacted to the recording. The whole of #7/1 is a reaction to the input, which starts with a

vociferous "What???" and goes on complaining that one of the speakers can barely be heard. This continues at the end of M2 with: "thank God for that!" (#7/2).

The first runs of protocols which did *not* begin with a fast progression were also examined to test whether the hypothesis could be disproved. Reactions to speech rate were found in two further instances for this recording: "That's gone a bit fast there" (#9/1); "Pretty quick" (#10/1), as well as one reaction to difficult perceptual conditions "I still can hear everybody! ... Aaaah! Shut up!" (#5/2).

- *Recording B1 (fat woman at doctor's):*

Three of the move logs began with a straight run (although everybody used fast progressions throughout the task). There was no evidence to support the hypothesis in any of the protocols. The only reaction found was an evaluative statement: "It's not actually that difficult" (#15/1), and this comment was not made immediately after listening, but at the end of the discussion that followed the move.

- *Recording F (forthcoming football match):*

Five of the 6 protocols for recording F began with a straight run: #17, #18, #19, #20 and #21. However no supporting evidence was found in any of them. The remaining protocol (#22) was a talk-aloud protocol from an individual subject, which also began with a fast progression (2M). In this case there was supporting evidence: "He says it quite quick doesn't he" (#22/2).

Another reaction was often observed in the first run, in the form of evaluative comments of the type already cited for recording B1: "It's not actually that difficult" (#15/1); "I understand the first few sentences" (#9/1). This initial assessment is negative in all the other cases recorded: "I don't really quite catch a lot of it" (#1/1), "Haven't got a clue what they're talking about" (#2/2), "I get the odd word" (#3/3; #9/3). Self-evaluation throughout the task will be examined in section 9.3.2.9.2 (p.236).

Summary

- Hypothesis 1a was verified for recording D (burglary): reactions to the speaker's delivery were observed in 4 out of 4 fast runs.
- It was not verified for recording C1 (interview): reactions to the speaker's delivery were only observed in 1 out of 2 fast runs; but they were *also* found in 3 out of 4 slow runs that were not considered in the hypothesis.

- It was not verified for recording B1 (fat woman at doctor's): no such reactions were observed.
- It was not verified for recording F (forthcoming football match): reactions observed in only 1 case out of the 6 fast runs.

Discussion

Immediate reactions to the input were found in recordings D and C1. However, the fact that they occurred in some protocols that did not display fast opening runs (3 cases for recording C1) suggests that the strategy is characteristic of the first run in a task, regardless of whether the first run in question is fast or slow.

In contrast, no cases of immediate reactions were found for recordings B1 and F. This may be due to the fact that these two recordings did not present the same initial difficulties for the subjects as the former recordings. In both cases, the context was immediately clear to them, since recording B1 began with a clearly enunciated statement setting the situation ("doctor, I have a problem"), and recording F provided a set of questions that activated the "football match" schema. Speech rate was also slower in recording B1, although not in F.

In the light of these results, the original hypothesis could be reformulated in terms of the initial reaction to "difficult" input, regardless of the pace followed by listeners in their first run: it appears that when input is "difficult", reactions to speaker's accent and delivery occur after the first run, but there are no such initial reactions when input is "easy".

9.3.2.1.2 Schema activation

Results

- *Recording D (burglary):*

Overt attempts to activate relevant schemata were found in all four cases "Maybe he was being interviewed" (#1/1); "He's going on about his house quite a bit" (#2/2); "Is he talking about the people in the house?" (#3/3) "Is he describing their jobs?" (#4/1);

- *Recording C1 (interviewing a TV presenter):*

One overt attempt was found: "They were talking about ... what he does ... where he lives" (#7/2). There is also some evidence in #9, where in the very first move one of the subjects mentions that the speaker "started talking about what he worked as" (#9/1). However this occurs within a fairly slow progression (run I = 7M) which does not fit the hypothesis.

In the rest of recording C1 protocols, bottom-up decoding began straight away and any schema activation that may have taken place as a result of decoded items was not directly observable.

- *Recording B1 (fat woman at doctor's):*

Only one instance was found, in protocol #14: "She goes to the doctor, she's got a headache" (#14/3). In all the other cases there was no demonstrable evidence that subjects were consciously attempting schema activation.

- *Recording F (forthcoming football match):*

No demonstrable evidence was found here either. Furthermore, the items that were decoded right from the beginning tended to be related to detail (the teams playing, the weather forecast, etc.) rather than top-level gist information (the speaker is a football correspondent).

Discussion

When input is easy to process, for instance in the case of recording B1 (fat woman at doctor's), schema activation strategies cannot be observed because activation is automatic: only one pair (the least successful subjects in the group) was found to use overt schema activation in recording B1. As seen in section 9.3.2.4.2, the rest of the subjects began to process entire chunks automatically as soon as they heard the recording (often on-line as it was still playing). In such cases, schema activation would have been triggered automatically by information readily available through decoding processes, thus avoiding the need for overt schema-activation strategies. If the opening sentence: "Doctor I have a problem" is immediately decoded as a whole chunk, the schemata "doctor's surgery" and "problem-solution" will not require further negotiation in order to become active.

In the case of recording F, the "football match" schema was already activated by the questions in the task brief, so even the weakest subjects were able to start filling in the relevant information slots (trying to answer the questions) straight away.

In contrast, recordings D and C1 did require a conscious effort in order to work out which schemata might apply. This was obvious for recording D (burglary), where all pairs made explicit attempts to activate one schema or another in their first run. This also happened in two cases for recording C1 (interviewing a TV presenter), although the fact that one of them used a slow progression suggests that the hypothesis is true of any run used at the beginning of the task for this type of recording, not only of first runs that use fast progressions.

The lack of observable schema activation in the opening runs of the remaining 4 cases for recording C1 needs to be accounted for. The explanation could lie in the way in which information is structured in this recording. The fact that it is an interview in which the speaker gives his personal details is obvious from the very beginning: "me llamo" (="my name is") is after all one of the first phrases foreign students learn on any beginners' course. The interview format is also obvious from the structure of turn-taking in the recording. Therefore, decoding of the initial segments can begin straight away (although not automatically, as was the case in recording B1). The "occupation" schema is also relatively easy to identify, even if the relevant information slots are difficult to fill in. Real difficulties in finding a suitable schema arise only later in the recording, after the question "Where are you from?", which is followed up by a highly abstract question on a topic unfamiliar to the students: "How is a person from Madrid perceived here in Motril?". Attempts to activate a relevant schema for this section of the recording were made at a later stage if at all. Subjects preferred to focus first on the initial segments which they were partly able to decode and could (sometimes automatically) relate to the schema: "personal details in interview"

The occurrence of overt schema-activation strategies in a slow first run (#9/1) for recording C1 suggests once again that this strategy is characteristic of first runs under certain conditions, whether or not the first run in question is fast or slow. Overt schema-activation strategies occur when subjects cannot activate a suitable schema by the time their first exposure to the input has taken place. The pace of such first exposure does not appear to be relevant.

9.3.2.1.3 Locating relevant information

Results

- *Recording D (burglary):*

None of the protocols for recording D showed any explicit evidence that the subjects were locating relevant information.

- *Recording C1 (interview):*

Two explicit cases were found, but both occurred within slow progressions which did not fit the hypothesis: "What's his name? ... Let's carry on with it and we'll sort that back." [sic] (#5/2, in run I =5M); and in #9, where one of the subjects instructed his partner to rewind the tape to a particular section "until he started talking about what he worked as" (#9/1 in run I =7M).

- *Recording B1 (fat woman at doctor's):*

None of the protocols for this recording showed any demonstrable evidence that the subjects were explicitly "locating" relevant information. It was found that in 5 out of 6 protocols, entire chunks were decoded straight away and uttered directly in L1, either while the tape was still playing (#11/1, 12/2, 13/1, 16/1) or immediately after the first move (#15/1).

- *Recording F (football match):*

The clearest evidence was found in 3 out of 6 protocols for this recording: #17 begins with a student reading out the question: "At the beginning of the recording a figure is given..." and proposing a hypothesis: "A price? Is it some pesetas?" (#17/1). "We definitely have some weather there" (#20/1); "There was a large number mentioned so it was either a price in pesetas or possibly the amount of people that are turning up for the match" #22/2). The other 3 cases are less obvious, yet the fact that in all of them, one of the test questions is answered straight away suggests that the subjects have successfully located the information related to the question: "Well I got the first one!" (#21/1); "He said something about 'el sol' and he said 'frío'... it's very cold and sunny" (#18/1); "Real madrid ... Barcelona" (#19/1).

Discussion

Given that recording F (football match) was accompanied by a set of questions requesting specific information, it is not surprising that the subjects attempted to locate information in the recording while they listened to it for the first time. This was

not necessary in recording B1 (fat woman at doctor's), since sufficient information was readily available through decoding processes straight away.

A similar effect was found for quite the opposite reason in recording D (burglary): since there was so little information that the subjects were able to decode straight away, they failed to pinpoint any particular section of the recording as potentially relevant. They were still too busy trying to activate a workable schema on the basis of the few words that they had managed to recognise.

In recording C1 (interviewing a TV presenter), the two cases who located relevant information during their first run did so on the basis of the two easily identifiable schemata that have already been discussed: "personal details" ("What's his name?... let's carry on and we'll sort that back" [sic]- #5/2) and "occupation" (instructing partner to rewind the tape "until he started talking about what he worked as" - #9/1).

Again, this strategy is not exclusively found in cases of fast opening progressions, since the two examples found in recording C1 actually used slow first runs.

It appears from this discussion that the strategy of "locating information that sounds relevant" is used at the beginning of a task, *regardless* of the pace of progression used in the first run, and provided that the following conditions are both met:

1. a suitable schema has been identified
2. the specific information required in order to fill in the slots in that schema has not yet been decoded

This would explain why the strategy was not found in recording D (condition 1 had not been met) or in recording B1 (condition 2 did not apply).

9.3.2.1.4 Summary of results

The table below summarises the results presented in this section.

Hypotheses on strategies used	*Recording D N=4 (all fast starts)*	*Recording C1 N=6 (2 fast starts)*	*Recording B1 N=6 (all fast)*	*Recording F N=6 (all fast starts)*
Getting acquainted with speaker's accent & delivery	**Confirmed** (found in all cases)	**Not confirmed** (found in 1 fast start and in 3 slow starts)	**Not confirmed** (not found)	**Not confirmed** (not found)
Activating relevant schemata	**Confirmed** (found in all cases)	**Not confirmed** (found in only 1 fast start and in 1 slow start)	**Not confirmed** (found in only 1 case: #14)	**Not confirmed** (not found: directly processing detail)
Locating info that "sounds" important	**Not confirmed** (not found)	**Not confirmed** (found in 2 slow starts only)	**Not confirmed** (not found)	**Confirmed** (found in all cases)

Table 9-5: Use of a fast run at the beginning of a task

9.3.2.2 Fast progression at the end of a task

9.3.2.2.1 Results

Hypotheses related to the last run of a PL task could only be tested with protocols in which the task had been completed, as was the case with all protocols from recordings B1 and F.

With recordings D and C1, some pairs went on working until time ran out, but had already reached a comprehensive meaning hypothesis for the recording by the time this happened. In such cases the task was regarded as "completed". Cases in which the pairs were were still engaged in intensive meaning negotiation when time ran out were not considered for this hypothesis.

The final runs of "completed" protocols were examined for evidence of hypothesis-checking strategies, with the following results.

- *Recording D (burglary):*

The task was completed in 2 out of 4 cases: #1 and #4, both of them ending with a fast progression.

The final run of #1 is centered on establishing the meaning of the word "carrera", with two hypotheses proposed: "retirement" and "graduation". There is clear evidence of hypothesis-testing at the end of this run: "She gave us a word for retiring and he would have said if it were retirement"; "definitely celebrating anyway" (#1/6). One last check of the related segments is made in M7 (ss.1-3).

There is also clear evidence in the other case: "I didn't recognise anything that could have been stolen" (#4/8). In both cases, evidence is in the form of comments made once the tape has stopped, and the strategy used is to seek evidence for an inference in the form of words related to the relevant schema.

- *Recording C1:*

Only in one case was the task completed for this recording (#9). It is also the only case where the last run recorded is fast (2M). The task is deliberately terminated by one of the partners after the second run, in which one hypothesis is tested at word-recognition level for an L2 chunk: "There we are: 'Soy de Madrid' " (#9/9).

- *Recording B1:*

The task was completed in all 6 cases, but only 4 will be examined here, since #11 and #12 were completed in a single run.

The first relevant case is #13, in which there are several instances of refining hypotheses at decoding level: e.g. "problem" (#13/1) becomes "big problem" (#13/2), "eating between meals" becomes "she says she's not gonna be able to not eat between meals", and one hypothesis rejection "something for tomorrow" becomes "What did you eat this morning?". These are all concentrated on the first half of the recording (up to s22).

In protocol #15, the last run (run III) is incomplete, but the task was possibly regarded as completed after run II: "That's it.... Yeah, just once more from the beginning" (#16/4). Run II included one global recap in M2, and several confirmations of decoded items in L1 throughout M2 and M3: "big problem", "she likes swimming" etc., one of which was refined from "calories" (#15/1) into "many calories" (#13/3).

Protocol #16 (produced by a subject working on his own), included numerous repetitions of decoded items while the tape was playing (#16/3-5), many of which were refined in the process.

In one case (#14), there was no evidence of hypothesis-checking strategies, as one of the subjects was still trying out decoding hypotheses at word level

("pelo"/"puelo"/"puede") when the other decided unilaterally to terminate the task (#14/5).

- *Recording F:*

All 6 protocols will be examined here as the task was completed in all cases. Three ended the task with fast progressions (#19, #21 and #22) and three with slow progressions (#17,18 and #20).

All three cases using a fast progression at the end of the task did so in order to check the answer to one particular test question related to the weather forecast for the match: "Good cool conditions" (#19/10); "I've actually got one minute or two, so I'll listen to it once more and see if I can get the weather forecast for the match" (#22/12). In one case, an intention to check through all work previously done is stated as well: "Shall we just listen to it again, just to get that? [the weather forecast] ... and then just double-check as you listen to it" (#21/9).

The pairs using slow progressions showed a similar focus on one particular question (the time of the match): "I don't know, tomorrow is 'pasada' or 'próxima' ... oh no it's not" (#17/8); "Yeah. -Tomorrow" (#18/17).

There is only one pair who did not use this strategy: in protocol #20, hypothesis-checking is an on-going process throughout the very long second/final run (27M). Hypotheses are tested one at a time as the successive questions are addressed sequentially in clusters.

9.3.2.2.2 Discussion

Table 9-6 summarises the results presented in the previous section.

All the pairs who completed the task used their last run to check previously made hypotheses. The only exception to this is pair #14 (recording B1), who terminated the task without evidence of any hypothesis-checking. Once the most assertive student in this pair formulated a situational hypothesis ("she goes to the doctor, she's got a headache" - #14/3), he assumed that the task had been completed successfully ("I'm done" #14/3). Hypothesis-checking did not appear to be seen as necessary by the student in this instance.

Strategy	***Recording D*** *2 completed (both end fast)*	***Recording C1*** *1 completed (ends fast)*	***Recording B1*** *4 considered (all end fast)*	***Recording F*** *6 completed (3 fast, 3 slow)*
Hypothesis checking at the end when the final run is fast	**Confirmed** (found in all cases)	**Confirmed** (found in the only relevant case)	**Confirmed*** *With one exception: #14.	**Not confirmed** (found in all 3 "fast" cases, but see below)
Hypothesis checking at the end whatever the pace of progression	No cases ended the task with slow progression	No cases ended the task with slow progression	No cases ended the task with slow progression	**Confirmed** (found in 6 cases, both fast and slow)

Table 9-6: Use of a fast run at the end of a task

In all the other cases in which the task had been completed, hypothesis-checking was used in the final run, albeit at various different levels, ranging from global hypotheses that sought to account for the meaning of the entire recording (#4/8), local hypotheses in the form of single words heard in the L2 ("zapatos" #2/21), or the final formulation for the answer to a particular test question (#19/10). This final stage was also used in order to refine previously made hypotheses (#16/3-5).

Once again, the presence of hypothesis-checking strategies in two of the cases that used slow final runs in recording F seems to indicate that the strategy is used at the end of a task, regardless of whether the final run is fast or slow. Since the other 3 recordings did not produce any cases of completed tasks using fast final progressions, it is not possible to ascertain whether the same strategy would have been used with these recordings, had the final run been slow. The evidence found so far suggests that hypothesis-checking is a typical strategy in the final run of a listening task, regardless of the pace of the final run in question.

9.3.2.3 Top-heavy patterns

9.3.2.3.1 Results

All move logs for recording D displayed top-heavy playback patterns, whereas there were no cases for recording C1, one for recording B1 (#14), and two for recording F (#19, #22). The relevant protocols were examined for evidence of two hypothetical processes:

1. *Hypothesis 5.a:* Identifying a familiar formal schema (e.g. problem-solution) and trying to fill in the slots for the first component of such a pattern.
2. *Hypothesis 5.b:* Identifying referrents that are used in subsequent referring expressions and anaphoric reference.

- *Recording D (burglary):*

Protocols from this recording were examined in detail as all cases shared a similar top-heavy pattern.

A number of words from the initial segments of the recording were recognised (and processed to varying degrees) by all or most pairs at the beginning of the task. Table 9-7 below shows the move numbers in which word recognition was first recorded for each pair:

	amigos (friends)	*médico* (medic)	*casa* (house)	*celebrando* (celebrating)	*el fin de carrera* (end of degree)
Pair #1	M1	M1	M2	M5	M3
Pair #2	M2	M4	M2	M7	M6
Pair #3	M2	M4	M3	----	M2
Pair #4	M1	M2	M1	M3	----

Table 9-7: Most recognised words in recording D

On the basis of these recognised items, all pairs attempted a number of inferences in order to establish a hypothetical schema of some kind or another: "being interviewed" (#1/1) "describing his house" (#1/2), "going on about what his friends are doing. One of them is a doctor" (#2/4), "going through the house, describing people in different rooms" (#3/9). None of the protocols cited so far (#1-3) show any evidence of subjects successfully recognising the formal schema as a narration in which the initial segments establish a setting for the action itself, that occurs later. In two cases this second part is overtly ignored: pair #1 ended the task with a statement of their final meaning hypothesis, which only incorporated information from the first 3 segments of the recording: "It's the end of the career of one of his friends, so they're celebrating" (#1/7). Although reasonably accurate in terms of propositional meaning, this final hypothesis shows no attempt to account for any elements from the rest of the

recording. As for pair #2, they embark in a time reference in the future tense: "and he's gonna do this, and he's gonna do that" (#2/18), which is clearly incompatible with the narrative schema of the recording.

Only one pair (#4) successfully recognised the "setting-action" schema. This however did not take place until the end of the task. Initially, they formulated similar hypotheses to those of the other pairs: "describing his friends' jobs" (#4/1), "he had them round his house" (#4/3), describing his "daily routine" (#4/4). But once "police" and "glass on the floor" were recognised from later segments in the recording (#4/5-6), the pair combined them together into a key inference "he's been burgled" (#4/7). Consequently, the initial segments of the recording were re-interpreted retrospectively as the (relatively unimportant) setting for the "burglary" story: "he could be at the doctor's or something , and he came back and found his house had been burgled" (#4/7).

In a few instances, subjects focused on decoding non-lexical items such as prepositions, tense morphemes, determinants and adverbs related to the lexical items that have been listed above: "esta amiga" > "Not 'friend', *this* friend" (#3/10); "con os amigos" [sic] > "with my friends" (#2/6); "como médico" (#4/3); "Is it his house?" (#4/8); "estábamos"(#3/8) >"they were in the house" (#3/9). In this last example the strategy led the subject to identify the use of the past imperfect tense. This detail could have given a clue as to the status of this part of the recording (i.e. the "setting" for the rest of the story). However the pair failed to make the necessary inference.

In summary, all pairs used the initial segments in order to build meaning hypotheses on the basis of a schema of some kind, but there was no evidence that such a schema took the form of a particular formal schema. Therefore hypothesis 5a was not confirmed for recording D.

Attempts to decode non-lexical items in order to establish the precise status of previously decoded lexical items could indicate efforts to identify accurate referrents for subsequent referring expressions in the recording. If this interpretation is accepted, it would constitute supporting evidence for hypothesis 5b .

- *Recording B1 (fat woman at the doctor's):*

Only one pair (#14) used a top-heavy pattern for this recording. The situation of going to the doctor's (#14/3) did trigger the "problem-solution" schema: *"she's got a headache"* (sic.) (#14/3), "doctor recommended swimming" (#14/5). As no effort was made to confirm the actual nature of the problem, this led to serious difficulties with

the rest of the recording. Nevertheless, hypothesis 5a was confirmed for recording B1. No supporting evidence however was found for hypothesis 5b.

- *Recording F (forthcoming football match):*

Two move logs (#19 and #22) displayed top-heavy patterns for recording F.

Pair #19 proceeded almost exclusively in straight runs, and the only move that was not a full straight run was M7, covering the opening segments (ss.1-7). Only moves 2-7 showed evidence of processing information contained in those initial segments, namely the price for the match tickets. Move 1 concentrated on the teams playing, and moves 8-9 on the weather forecast for the match. An almost identical processing pattern was found in protocol #22, even though the move log looked very different from the previous pair's: move 1 focused on the teams playing, moves 2-6 on the price for the tickets, and the rest (moves 8-13) on the weather forecast.

These findings suggest that the subjects were seeking specific information related to the questions provided in their task brief, rather than looking for key referrents or discourse schemata. Hypotheses 5a and 5b were not confirmed for recording F.

9.3.2.3.2 Summary and discussion

Table 9-8 below summarises the results presented in this section.

Strategy	***Recording D*** *N=4* *(4 top-heavy)*	***Recording C1*** *N=6 (no top-heavy cases)*	***Recording B1*** *N=6* *(1 top-heavy)*	***Recording F*** *N=6* *(2 top-heavy)*
Identifying a specific formal schema	**Not confirmed** (looking for a schema, but not a *formal* schema)	No cases of top-heavy patterns	**Confirmed*** (*note that there is only 1 top-heavy case: #14)	**Not confirmed** (seeking specific info related to the questions)
Identifying referrents	**Confirmed** (found in 3 cases out of 4)	No cases of top-heavy patterns	**Not confirmed** (no evidence found)	**Not confirmed** (as above: no evidence found)

Table 9-8: Use of top-heavy patterns

Two hypothetical explanations had been proposed to account for the top-heavy pattern phenomenon:

1. Subjects try to identify specific formal schemata (e.g. problem-solution) from the initial segments of a recording
2. Subjects try to identify the referrents that are to be used in later referring expressions throughout the recording

Here again, any evidence found in the protocols would indicate that these processes are taking place as controlled strategies rather than as automatic processes (given that the latter would result in an absence of observable evidence).

When they start working through a recording in L2, listeners first need to achieve a minimum level of schema activation in order to account for one (and ideally more) of the following aspects:

- the situational schema of the message (e.g. "at the doctor's")
- the formal schema followed in the message (e.g. "problem-solution")
- the topic of the message (e.g. "being overweight")

From that point (but only then) can listeners start using any clues available to help fill in the information slots assumed to be contained in the active schemata.

If for any reason the necessary schemata are activated straight away, they can proceed to search for relevant information throughout the recording to fill in as many slots as possible.The resulting playback strategy is not top-heavy, as shown in the case of recordings B1 and F:

Not surprisingly, there was no evidence to suggest that the subjects were seeking a formal schema in *recording F (football match)*, given that the questions activated the relevant schema ("sports correspondent talking about forthcoming football match") and provided sufficient clues as to the information that was to be found in the recording.

In *recording B1 (fat woman at doctor's)*, the problem-solution pattern was immediately activated as soon as the subjects successfully decoded the utterance "Doctor, I have a problem" and no explicit processing was required. In some cases, this statement was not even uttered, and the pair directly went on to report the nature of the problem in question: "she eats too much". The five cases who displayed this behaviour did not follow top-down playback patterns. The only pair who produced observable evidence of formal schema activation was pair #14, who explicitly uttered the bridging inference

"she goes to the doctor" and invented a problem ("a headache") to complete the scenario.

When subjects experience difficulty finding suitable schemata, they try to decode as much information as possible from the beginning of the recording by means of controlled bottom-up processing. This produces the type of top-heavy playback patterns that were consistently found for recording D.

Recording D (burglary) was in striking contrast to the other recordings, as all four move logs displayed top-heavy patterns. There was evidence to support that all pairs were trying to generate schema hypotheses on the basis of a few decoded words from the beginning of the recording. However, most only managed (at best) to work out the first situation presented ("I was at a friend's house, celebrating her degree"). Nevertheless, they still failed to realise the status of this information within the overall formal schema (i.e. the fact that a "when suddenly..." section was about to come, in which the relevant event actually started). Protocol #1 is a clear example of this. Only one pair (#4) realised what the complete schema was, and this was not because of accurate decoding of the opening segments, but as the result of an inference based on the second half of the recording, in which they understood that the recording was about a burglary.

In their search for clues and when all else fails, subjects may even resort to controlled processing of non-lexical items (e.g. tense morphemes) related to the few lexical items that they have managed to decode. This strategy, applied to the opening segments of recording D, could be interpreted as a desperate attempt to establish the precise nature of referrents that have not yet been fully identified in the absence of further lexical clues.

The absence of top-heavy patterns in *recording C1 (interviewing a TV presenter)* needs to be explained. The schemata: "interview" + "personal details" + "occupation", were easily identified from the first few segments. However, this information was of little use for the global comprehension of the recording. The key to the gist was not likely to be found at the very beginning, but more probably in relation to the speaker's occupation itself. Therefore four pairs produced high-RR clusters somewhere around segments 6-11, in which the speaker first states his occupation.

There could, however, be a much simpler explanation for top-heavy patterns: when they do not know very well what information a particular section of the recording may contain, subjects might fear that if they just rewind the tape a little, they could "land" half way through an idea, and thus fail to understand what the section is about. By rewinding right to the beginning, they stand a better chance to ascertain the status of that information within the global structure of the recording. In contrast, when subjects have a pretty good idea of the type of information that they are looking for, they can afford to skip the beginning and go directly where they know that the information will be found. Since they know what they are looking for, they can easily ignore any irrelevant segments that may have been played back accidentally in the same move.

A top-heavy pattern is bound to appear if listeners systematically rewind the tape right back every time they want to hear again a passage that causes them difficulty anywhere in the recording. Starting afresh from the beginning is a natural human strategy in any skill area, whenever too many problems accumulate while performing a task. This interpretation would explain the existence of an overall tendency to play the beginning of the recordings more than the end (even in cases not strictly covered in our definition of "top-heavy" patterns), and why the frequency of playback seems to decrease gradually through the recordings in most cases. The tendency was less evident for recording F (football match) than for recordings D (burglary) and C1 (interview), since the questions provided with F guided listeners directly to the relevant sections in the recording. Recording B1 (fat woman at doctor's) was the only one that did not produce move logs showing such a tendency at all. This could be due to the fact that the message was largely decoded through automatic processing. The need to rewind right to the beginning before the end of a run is more likely to arise when the flow of processing is significantly delayed by a problematic segment. If progression through the task is smooth, there is no need to start afresh in the middle of a run.

9.3.2.4 Playback sequence in recording F

9.3.2.4.1 Results

Hypothesis 7 proposes that when tackling a recording in which the information follows a different sequence from the sequence of questions in the task brief, subjects follow the sequence of questions rather than the sequence of information in the text.

Recording F (football match) was the only one that was accompanied by questions in the task brief. However, the sequence in which such questions were presented did not match exactly the chronological sequence in which the information related to the questions was mentioned in the recording. This feature allowed hypothesis 7 to be tested by examining which of the two sequences the subjects had followed in order to complete the task.

Table 9-9 below illustrates the discrepancies between the two sequences in recording F:

Sequence of information in the recording	*Sequence of questions in the task brief*
• Time of broadcast • Figure given • Teams playing (1st mention) • General mood ("passion", "league match") • Time of match / Weather forecast • General mood ("spectacular") • Teams playing (2nd mention)	• Question 1 - Teams playing • Question 2 - Time of match • Question 3 - Weather forecast • Question 4 - Figure given • Question 5 - General mood

Table 9-9: Information sequence in recording F

According to this, one could logically infer that a question-driven sequence has been followed when:

(a) The teams playing are the first information mentioned in the protocol

(b) The two items "time of the match" and "weather forecast" (which are mentioned together in the recording) appear in distinct sequence in the protocol: first the time, then the weather.

(c) Information mentioning "passion" and "league" (in relation to the general mood in the city) appears only at the very end of the task.

Conversely, a text-driven sequence has been followed when:

(d) Information related to the "figure given" in the recording clearly appears before information related to the "time of the match" and "weather forecast".

Table 9-10 summarises the occurrence of these features in the protocols for recording F. Positive entries indicate cases in which the feature was clearly evident. Any ambiguity was recorded as a negative entry.

	Question-driven sequence			***Text-driven***
Protocol	***(a)*** *"Teams" = 1st item mentioned*	***(b)*** *"Time" clearly before "weather"*	***(c)*** *"Passion /'league" = last mentioned*	***(d)*** *"Figure" before "time" & "weather"*
#17	N/A ("teams" not mentioned at all)	No	No	Yes
#18	N/A ("teams" not mentioned at all)	No	No	No
#19	Yes	No	No	No
#20	Yes	Yes	Yes	Yes
#21	Yes	No	Yes	Yes
#22	Yes	No	Yes	Yes

Table 9-10: Question-driven and text-driven sequence features in recording F

A glance at the table shows that there were no cases fully consistent with either hypothesis, as protocols included a mix of features from both question-driven and text-driven sequences.

Two protocols did not make any mention of the teams playing. The information was understood, since "Real Madrid" and "Barcelona" were correctly entered in both pairs' answer sheet, but it was impossible to establish at which point that information had been obtained. The lack of discussion and the reactions in the rest of the group would suggest that the item did not cause any difficulty as all the other pairs reported the answer in their very first move.

Cases #19 and #20 illustrate two different approaches to the task:

Pair #19 proceeded almost exclusively in straight runs, gleaning as much as possible from all parts of the recording each time, elaborating and filling new gaps each time, as they worked on all the questions in parallel. The sequence of responses seemed dictated by the subjects' own incidental sequence of processing, and the process continued until all questions were eventually answered.

In contrast, the playback sequence used by pair #20 was very closely determined by the sequence of questions, and the subjects explicitly stated their awareness of sequence discrepancies between questions and text in every instance: question 4: "It says 'at the beginning of the clip', so do you think that [this question] should be there?" (#20/3); question 1: "you think it starts at number one, but then right at the end it says Real Madrid and Barcelona" (#20/3); question 5: "now there's a bit before but there's a bit at the end as well" (#20/25). In terms of strategy, they tackled the questions in strict sequence, with the exception of question 4 ("figure given"), which was restored to its most logical location at the beginning of the recording. There was clear evidence of question-driven strategy throughout the task: "We can go on to something else then. All right what's this? The weather?" (#20/14); "Right what's the last one? 'General mood'." (#20/25).

The logical assumption that subjects had the questions in mind while listening to the recording was supported in all cases by the relevance of the information summarised after the first move, and by comments like "Well I got the first one!" (#21/1) made immediately after the tape was stopped.

The student who was working on his own explained the strategy in great detail as follows: "I listen to it all the way through and I pick up on a few words here and there (or as much as I can, obviously)... and then I go back to the individual points with respect to the questions .. and in some way I have to filter out what I don't need. And then obviously I pick up on a word, realise that it's to do with the question, and then try to pad around that word, to see the other words he uses, so I can actually make sense of it." (#22/7)[30].

In summary, hypothesis 7 was not confirmed, since none of the protocols showed a purely question-driven or text-driven playback sequence. Undoubtedly questions did have an effect on task sequencing, but this was of a more complex nature than anticipated. Observed playback strategies appeared to combine sequencing factors related to the questions, the text, and also the subject's natural sequence of perception and processing (see discussion below).

[30] Note the difference between the type of information supplied by this subject (who was working on his own while introspecting) and the kind of information obtained from paired listening protocols (all the other material quoted in this section). In this particular instance, introspection confirms observation, but the two methods clearly operate in very different ways.

9.3.2.4.2 Discussion

It was clear from the results that neither a purely question-driven or text-driven sequence was followed by any of the pairs. What was found was a range of strategies that different pairs combined in different ways.

The initial steps appeared to be common to all the cases examined:

Step 1: read the questions

Step 2: listen to the recording straight through (twice in the case of #18), with attention to any information relevant to the questions

Two facts confirm the use of this strategy by all subjects: all move logs from pairs working with recording F begin with a straight run; and all the responses recorded after that initial straight run comprised information that was directly relevant to one or more questions from the task brief.

The outcome of this step is twofold:

- Some information is processed straight away, ready to be utilised in response to the relevant question (e.g. "Real Madrid" and "Barcelona" for the "teams playing").
- Some of the output is in the form of partly processed information, which is used to locate the part of the recording in which the answer to a particular question can be found (e.g. "we definitely have some weather in there" #20/1). All cases displayed this strategy as a result of their first move (see p. 197).

The first type of output (fully processed information) generates a similar response across the group:

Step 3: utilise the information that is fully processed (i.e. answer the relevant question from the answer sheet).

Evidence of this behaviour was significant: 4 protocols out of 6 gave the names of the two teams straight away. The remaining two failed to mention them explicitly at any point during the task, but entered them correctly on the answer sheet. This indicates that the information was obtained automatically and did not require any amount of meaning negotiation in these two cases.

Step 4: fill in the gaps.

It is at this stage that the strategies found begin to differ across the sample:

- *Strategy A:*

  Go back to the information that has been "located" and process it. In cases where relevant information has been located for more than one question, subjects begin with the first item that occurs in the recording (#21/1-2).

- *Strategy B:*

  Save "located" information for later and try to locate the section that is relevant to the next question listed. This strategy was adopted through most of the task by pair #20.

- *Strategy C:*

  Repeat step 2 above (straight run), processing whatever you can, with particular attention to information that has already been located. This strategy was used by pair #19, who proceeded almost entirely in straight runs, processing several questions in parallel, until the only gap left was the "weather forecast", on which the last 3 straight runs were spent.

Often protocols present a combination of the strategies described above, making it impossible to find a "pure" strategic pattern. Two more factors seem to play a part in playback strategy decisions and ought to be considered in addition to "sequence of questions" and "sequence of information in the text". These are: the saliency of certain items in the recording, and the level of complexity of the information required to obtain the complete answer to a given question.

- *Saliency:*

The most salient items will be located faster by the subjects. For instance "el sol" [=the sun] was salient for two reasons: it was pronounced with a stress in the speaker's utterance and followed by a pause; and it was a very familar word to the students. Consequently, several pairs located the "weather" section straight away. In one of their experiments on controlled and automatic processing, Shiffrin and Schneider found that, when subjects come across an item that they have learnt to detect automatically through previous training, their attention automatically (and unstoppably) focuses on it, even when they were supposed to be engaged in controlled search for a different item present in the input (Shiffrin and Schneider 1977:187, reporting the results of experiments 4a-4d). Here, subjects who were looking for the answer to "time of the match" (=controlled search) might have been distracted when they heard and recognised the word "sol" (=automatic detection), and

consequently shifted the focus of their search from "time of the match" to "weather forecast" (see also Rabinowitz and Chi 1987).

- *Complexity of information:*

Even salient information may necessitate further processing if the answer requires a combination of information items which are inter-related in complex ways (e.g. "The sun is shining now, we expect good weather for tomorrow, a bit cold, but almost ideal" - from ss.16-23). In this recording for example, the names of the two teams playing are not only salient (a feature that makes them easier to locate), but also extremely simple to process (which means they can be utilised without further playback). In contrast, the weather forecast (easily located because of the saliency of "sol") will almost certainly require several moves before a complete answer can be reached.

9.3.2.5 Fast progression overall

9.3.2.5.1 Results

Two possible reasons were considered in order to account for fast overall progressions:

- *Hypothesis 3a* - Fast progression overall indicates that little processing is taking place.
- *Hypothesis 3b* - Fast progression overall indicates that most decoding processes are taking place automatically on-line.

The following move logs displayed fast overall progressions, either throughout the whole task or over a significant section of the recording. The results are summarised in Table 9-11 at the end of this section.

- *Recording D (burglary):*

No clear cases of overall fast progression through the whole task were found for recording D, where the average number of segments per move (s/M) for protocols #1, #2, #3 and #4 had values of 18, 9,16 and 18 respectively.

Move log #1 comprised a comparatively fast section in M6, covering s.7 to the end. Two limited attempts to decode segments located in the second part of the recording were found in this move, but they were inaccurate. Furthermore, at the end of the move the subjects did a global recap of all information gained, in which all information

reported belonged to ss.1-6. No information beyond this point was included in the final meaning hypothesis. The "no processing" hypothesis was further confirmed by the statement: "Even if I listened to it five times I wouldn't know. That's as much as I know." (#1/6)

- *Recording C1 (interviewing a TV presenter):*

Move log #8 showed an abnormally fast progression in M35 (s.36 to the end), which was also the only move in which the second half of the recording was ever played back. There was virtually no processing of the information contained in these segments. The only items stated in the protocol (#8/35) were: "local TV" (but "TV presenter" had already been decoded in M15): "casi" (with no follow up into L1); and "Mondril" (which was the only item possibly utilised in the next run to help decode an earlier segment and reach the correct form of "Telemotril" - #8/37). These three items represent a negligible proportion in comparison to the amount of processing that the pair did on the first half of the recording. Therefore the "no processing" hypothesis was confirmed in this case.

Move log #9 displayed a fast progression over the entire task, with an average of 28 s/M (well above the average for this recording, which was 12 s/M). The students' comments throughout the task speak for themselves: "I can only pick up the odd word you know, I totally lost gist of it" (#9/3); "I could listen to that about a hundred times, and I'm not gonna be able to make head or tail of it" (#9/5). One of the subjects even expressed his eagerness to end the task in several occasions: "Are we finished? Can we kill her [=the researcher], you know?" (#9/7); and when his partner tried to prompt a hypothesis ("I'm not quite sure why he wants to get to Madrid"), his reply was "I don't think we'll ever know" (#9/7); the task was finally terminated when the question "Shall we keep going?" was answered by a "Well yeah and no" (#9/9). The "no processing" hypothesis was confirmed in this case, with an additional element of "refusal to process".

- *Recording B1 (fat woman at doctor's):*

Fast overall progression was found to be a distinctive feature for recording B1 (see p. 187). In five cases out of 6, the protocols showed that large numbers of segments generated immediate responses in the form of chunks[31] uttered directly in L1, often immediately after they had been heard, while the tape was still playing: "She doesn't

[31] See p.129 for a definition of "chunk".

like [running], but she likes swimming" (#11/1); "I'm eating too much" (#12/2); "It has too many calories" (#13/1); "She can lose 10 kilos in a week" (#15/1); "What do you normally do in the morning?" (#16/1). In the first move alone, protocol #11 contained 5 such L1 chunks, #12 contained 8, #13 contained 9, #15 contained 6, and #16 was almost an instance of on-line simultaneous interpreting, where each utterance was translated immediately while the tape kept running. Therefore the "automatic processing" hypothesis was confirmed for 5 out of 6 protocols for recording B1.

The single exception was pair #14, which proceeded in fast progression as well, but

Recording	*Protocol*	*Section(s) where fast progression occurs*	*Confirmed hypothesis*
D	#1	M6 (s.3 to end)	No processing
C1	#8	M35 (s.3 to end)	No processing
	#9	Fast overall	No processing (+ refusal to process)
B1	#11 #12 #13 #15 #16	Fast overall	Processing automatised
	#14	Fast overall	No processing (+ refusal to process)
F	#19	Fast overall	No hypothesis confirmed (task-specific strategy?)

Table 9-11: Move logs displaying fast overall progression

did not display the high levels of automatised processing that were observed in the other five cases. Moves 1 and 2 were spent discussing mechanical matters of equipment operation, and the first signs of actual processing appeared in M3, where two statements were recorded: "She goes to the doctor's" and "she's got a headache" (#14/3). None of these would qualify as an L1 chunk, given that they do not correspond to the propositional meaning of the text, and must therefore be the result of inferences. In fact, the only item approaching an L1 chunk was "Doctor recommended don't do 'natación' " (#14/5), and the rest were all single words in

either L1 or L2 (but please see note[32]). As in the case of pair #9, one of the students is keen to end the task quickly: "That's it, shall we just... [partner recalls another word:"frío"]. Maybe, maybe. Shall we just fill the paper? OK we're done." (#14/5). As for pair #9, the confirmed hypothesis here is "no processing", with an added element of "refusal to process".

- *Recording F (forthcoming football match):*

Pair #19 displayed a distinctive strategy within the group, proceeding through recording F almost exclusively in straight runs. The type of processing taking place can be seen by examining a complete single move as an example. This is the conversation that took place after the straight run made in M3:

— It takes place at night.
— At night? He said something about tomorrow.
— Tomorrow night. Er... is it eighty for "ochenta mil pesetas".
— "Ocho mil", is that for a ticket?
— I don't know, he says it... could well be. Er...
— And he said er... it's like looking forward to a spectacular match. Is that right?
— A spectacle yeah. (#19/3)

We can see that processing is not entirely automatised ("at night" and "tomorrow" are further processed in M4). Eventually though, all the information required to answer the questions was successfully decoded, suggesting that neither hypothesis ("no processing" or "automatised processing") was confirmed in this case. Instead, the pair seemed to be processing several test questions simultaneously ("time of the match", "figure given" and "weather forecast" in the example). This strategy was discussed in more detail in section 9.3.2.4.2 (p.214).

9.3.2.5.2 Summary and discussion

The results can be summarised as shown in Table 9-11.

The processing strategies underlying fast overall progressions were found to be closely dependent on the nature of the input. Three different scenarios for fast overall progressions were identified:

1. Subjects process the input automatically when input is "easy"

[32] It must be pointed out that both students in this pair were non-English native speakers (Turkish and Greek), therefore what should have been L1 and L2 were in fact L2 and L3 for them. The implications of this anomaly will be addressed in the discussion.

2. Subjects cannot process the input when input is "difficult"
3. Subjects will not process the input, whether input is perceived as "easy" (the "no need" approach) or "difficult" (the "no point" approach)

Scenario 1 applied for most pairs in the case of recording B1 (fat woman at doctor's), in which the situational context was immediately stated in the opening segments. The "doctor surgery" and "over-eating problem" schemata and their related vocabulary were familiar to the students and the speakers had clear, moderately paced delivery. Therefore the subjects were able to decode substantial amounts of input automatically.

In the case of recordings D (burglary) and C1 (interview), the reason for progressing fast was the opposite. For a start, the context of situation in these two recordings was not obvious straight away: the apparent "party" scenario that started recording D gave way to a "burglary" scenario that was difficult to recognise; as for recording C1, all that could be inferred about the actual topic of the interview was that it could have something to do with the speaker's occupation.

The reason for the fast progressions over the second half of these two recordings was that the subjects were unable to process such difficult input. In both recordings, the second half presented even more difficulty that the first half. In the case of recording D, one could assume that "burglary" (mentioned in the second half) is more difficult than "party" (mentioned in the first half). In the case of C1, we already noted the appearance of a highly abstract, unfamiliar topic half way through the recording: "How is a person from Madrid accepted in Motril?".

An alternative (or complementary) explanation for the fast progression found in the second half could be that these cases are instances of top-heavy playback patterns (see section 9.3.2.3.2 above).

Finally, a third scenario was found both in "easy" and "difficult" recordings, represented by two cases in which the subjects just seemed reluctant to involve themselves in processing the input to the level required. Findings from the analysis of social/affective strategies could help explain such behaviours. The two cases in question (#9 and #14) were the only two pairs who displayed lack of cooperation from one of their members. In the case of #9 (who was working on a "difficult" recording: C1), this attitude seemed due to negative experiences in previous class tests (#9/2) and the stress of having his performance recorded (#9/7). The other case was #14 (who was using an "easy" recording: B1), where the leading partner did not show

signs of frustration or hostility, but simply appeared to consider that his task had been completed as soon as a (very coarse) hypothesis was formulated to account for the context of situation. These behaviours were labelled respectively the "no point" approach and the "no need" approach.

9.3.2.6 Slow progression overall

9.3.2.6.1 Results

Hypothesis 4 claims that during a slow overall progression, bottom-up processing is more intensive than when a fast overall progression is used.

The overall progression was considered to be "slow" when the s/M value (number of segments per move) in the individual move log was smaller than the average s/M for the whole sample of move logs from the same recording. Recording B1 (fat woman at doctor's) was not considered, given that even the slowest progressions for this recording were too fast overall (group mean = 40.6 s/M) to be regarded as "slow" progressions (see.Table 9-1 p.188).

The following move logs from the remaining samples were found to be slow overall progressions:

Recording	*Avge. pace in the recording*	*Move logs with "slow" progressions*	*Avge. pace in this move log*
D (N=4)	13 s/M	#2	9 s/M
C1 (N=6)	12 s/M	#8	6 s/M
		#10	7 s/M
F (N=6)	13 s/M	#20	7 s/M

Table 9-12: Move logs displaying slow overall progression

For the three recordings considered, the amount of bottom-up processing was estimated for every move log as follows:

Words, phrases and chunks corresponding to the propositional content of the recording were encoded as L2w, L2p or L2c respectively if they were uttered by the subjects in L2; and L1w, L1p or L1c if they were uttered in L1[33]. Only the first

[33] Full details of the encoding procedure are given in section 6.2.2.2.1, p. 128.

occurrence of these items was counted, unless an item was later developed into a larger unit (e.g. a single word, later integrated into a phrase or a chunk). Each occurrence of any of these items was counted as a "decoding" token (DT). The values found are shown on Table 9-13.

Move logs that were identified as clear instances of either fast or slow overall progressions are shown under the "pace" label.

The means obtained for DT values in the three groups (recordings D, C1 and F) were compared using a one-way ANOVA test, which showed no significant differences across the three groups overall.

Recording	DT	Move log	Pace	DT
		#1		27
D		#2	Slow	32
Mean	***27***	#3		23
Std. Dev.	*3.74*	#4		26
		#5		47
C1		#6		37
		#7		29
		#8	Slow	82
Mean	***37***	#9	Fast	11
Std. Dev.	*25.71*	#10*	Slow*	16*
		#17		12
F		#18		25
		#19	Fast	21
		#20	Slow	35
Mean	***20***	#21		21
Std. Dev.	*8.86*	#22*		11*
Recordings D, C1 and F:				
			Mean	***28.44***
			Std. Dev.	*17.39*

Table 9-13: Decoding tokens in fast and slow overall progressions

* *Think-aloud data from students working alone*

Since experimental design could not determine in advance which pairs were to operate in fast or slow overall progressions, only 6 clear-cut cases (4 slow, 2 fast) were available for analysis. Such limited sample size makes comparisons rather adventurous. Furthermore, move log #10 was not produced by a pair, but by a single student talking aloud, even occasionally interacting with the researcher. It was therefore discarded, as the processes taking place were likely to be affected by different performing conditions. The three remaining cases of slow overall progressions (#2, #8 and #20) certainly suggest that more decoding takes place when progression is slow, given that all three protocols produced the highest DT scores within their group (32, 82 and 35 respectively), all of which were more than

one standard deviation above the mean. This appears to be consistent with the fact that one of the two cases identified as fast overall progressions (#9), produced a DT score of 11, which was the lowest for its group.

Although encouraging, these results were not considered to offer sufficient evidence to confirm the hypothesis, given the small size of the sample available for analysis.

9.3.2.6.2 Summary and discussion

The findings obtained from the analysis of move logs that displayed slow overall progressions were inconclusive. However, the general tendency emerging from this limited sample shows that more decoding tokens were found in slow progressions than in the other instances.

It must be pointed out that a reduced number of decoding tokens does not necessarily imply that less decoding has taken place. It could simply mean that whichever decoding processes were operating did so in a manner that was not evident in the protocol transcripts. This is consistent with the findings discussed in the previous section with regard to fast overall progressions in which decoding processes are automatic.

Of course, fewer decoding tokens could also mean that no processing is taking place, neither automatic or controlled. It is not surprising to find that pair #9 (whose "no point" approach resulted in fast progression) also produced fewer decoding tokens than any other pair in its group. The fact that the same feature was not found in all fast progressions could be due to the fact that not all fast progressions are fast for the same reason.

As for the high numbers of decoding tokens found in slow overall progressions, they would indicate that a larger amount of *controlled* decoding processing was taking place. Hence the need to work intensively on a small number of segments at a time, since controlled decoding proceeds more slowly than automatic decoding, often one small step at a time (this strategy can be seen at work in protocol #8, moves 22-29). If confirmed, these findings would support the interpretation of slow progressions proposed by Merlet and Gaonac'h (Merlet and Gaonac'h 1995).

9.3.2.7 The a-b-a pattern

9.3.2.7.1 Results

Hypothesis 6 claims that when an a-b-a pattern occurs, its three cycles correspond to the following strategies:

Stage 1 (Fast run at start):

- *Getting acquainted with the speaker's accent and delivery*
- *activating relevant schemata*
- *locating information that "sounds" important*

Stage 2 (slow run in the middle)

- *Bottom-up processing more intensive*

Stage 3 (fast run at end)

- *Hypothesis-checking*

Only one case of a clear a-b-a pattern was recorded within the sample's 22 move logs: pair #4 (recording D: burglary).

Stage 1: in the initial fast run there was evidence of two out of three hypothesised processes: the subjects reacted overtly to the speaker's accent and delivery: "What a relief it stopped!" (#4/1) and tried to activate relevant schemata: "Is he describing their jobs?" (#4/1), but there was no evidence of attempts to locate information that "sounded" important

Stage 2: the responses recorded between moves 2 and 4 were mostly in the form of single words uttered in L1 or L2, which supports the bottom-up processing hypothesis. However, the pair also began to make global inferences at the end of M4:

> — Do you reckon he's leaving the house or something? Putting his shoes on and opening the door and shutting the door and... opening the door, well he forgot something.
> — Well it could be his daily routine. Yeah? He's talked about being at the doctor's or something.
> — But he lives with his father and his friends [laughs]
> — I don't think he does.
> — Well if he's opening all these doors it will have to be a big house" (#4/4)

From that point onward, every new word that is decoded is rapidly matched with an inference that integrates it into the global interpretation of the recording: "trabaja en casa" > "Well he does housework as part of his daily routine" (#4/5); "glass on the

floor" > "perhaps he smashed the window of the door" (#4/6); "the police" > "He's been burgled" (#4/7).

Stage 3: All previously formulated hypotheses were grouped together in the global recap that immediately followed the final straight run (M8). There was also evidence of hypothesis-checking in this final run, made quite clear in the remark "I didn't recognise anything that could have been stolen" (#4/8), which showed that the pair were trying to find supporting evidence for their "burglary" hypothesis.

In this particular case, the a-b-a hypothesis was confirmed, but with the following amendments:

- No evidence of the "locating relevant information" strategy was found in stage 1.
- During stage 2, intensive bottom-up processing gave way to global inferencing, which was then applied to any items subsequently decoded throughout this stage.
- Stage 3 was used not only for checking previous hypotheses, but also to group them all into a final interpretation for the entire recording.

However, it would not be wise to generalise these findings, since they are based on a single case.

9.3.2.7.2 Discussion

The slight differences found between the hypothesised a-b-a playback strategy can be explained in terms of input-specific and task-specific variables.

The strategy of "locating relevant information" was absent, simply because it was a strategy difficult to apply on this particular recording (recording D - burglary). Since no obvious schema was given at the start, subjects were not ready to start locating relevant information,as they did not know what to look for. The use of top-heavy playback strategies across the group shows that, with this recording, their efforts were primarily concentrating on finding workable schemata. If a similar a-b-a strategy had been found for a different recording such as recording F (football match) instead of recording D, one might have been more likely to observe the strategy of "locating relevant information" in "stage one" of the a-b-a procedure. On the other hand, evidence of "schema activation" might in such a case have been missing, as the strategy would have been unnecessary because the schema was already given by the questions provided with recording F. This claim is obviously speculative, given that the only a-b-a playback strategy found in this sample occurred with recording D.

In the example found, inferencing strategies (not predicted in the hypothesis) occur in "stage two" for a similar reason: in this recording, input is too difficult to be understood through decoding processes alone. It is likely that no inferencing would have been found if the a-b-a pattern had occurred using an easier recording such as B1 (fat woman at doctor's). In such a case, it is plausible to assume that no inferences would have been required, given that much of the input would have been decoded automatically. In contrast, comprehending recording D required that every single decoded item was explicitly utilised for hypothesis-generation. Hence the emergence of inferencing as soon as there were enough decoded items that could be integrated in order to form a global meaning-hypothesis.

The global recap found at the end of stage three is a continuation of the same strategy. The more steps it has taken to put together a global meaning hypothesis, the more need there is to sum up the results of the negotiation process and recap the final hypothesis at the end of the task. Had the recording been easier to process, the outcome of the task might not even have been stated explicitly, and any checks could have taken place on-line during the final run.

The specific patterns discussed above are clearly speculative, but the fundamental point of this argument remains the following: whenever the a-b-a strategy is used, the processes and strategies found in each of the three stages are largely the same as those proposed in the hypothesis. However, they may be affected by input-specific and task-specific factors, such the as availability of global schemata, the amount of input that can be decoded automatically, the availability of clues in the form of questions, and possibly a number of other factors not yet identified.

It is also worth noting that only one instance of the strategy was found out of 22 protocols, which seems to suggest that the strategy is not a typical behaviour. Nevertheless, it was used by the only pair who eventually managed to recognise the "burglary" scenario in recording D: this could imply that the a-b-a pattern is a highly effective playback strategy, provided that it is used in combination with effective processing strategies (such as the gradual introduction of inferencing in this instance).

9.3.2.8 Playback frequency and recall

9.3.2.8.1 Results

Hypothesis 8 was intended to account for the fact that the best recalled segments are also the ones that are played most often.

In order to test the proposed explanations for this phenomenon, the protocols were analysed with particular attention to the subjects' responses during the moves in which high-RR[34] segments were played back. The analysis focused only on the responses that were directly related to information contained in the segments in question. Any information related to adjacent segments was disregarded. The criteria and RR tables described in section 9.3.1.2 (p.189) were used in order to identify which segments should be regarded as high-RR segments.

Since there was significant evidence to support the existence of a fairly widespread tendency to play back the beginning of a recording more than the end, only the segments contained in the second half of the recordings were considered. This precaution was taken to prevent any possible interference from top-heavy pattern phenomena.

The analysis examined each of the high-RR clusters separately, searching for the following evidence:

(a) Cases in which new items are decoded when the segments are repeated.

(b) Cases in which further inferences are made when the segments are repeated.

(c) Cases in which word recognition occurs right at the beginning of the repeating sequence.

(d) Any combination of the above.

Each of these types of evidence would respectively support one of the hypothesised explanations for the frequency/recall phenomenon:

(a) Recall is better because greater exposure to input provides more opportunity for bottom-up processing. If this is true, Experiment 3 will show that the gain observed from one time to the next is mostly related to decoding strategies.

[34] As explained on p. 189, the repeat-rate (RR) of a segment is the total number of times it was played back by a particular pair of subjects during the task.

(b) Recall is better because increased processing time provides more opportunity for top-down processing. If this is true, Experiment 3 will show that the gain observed from one time to the next is mostly related to inferencing strategies.

(c) Recall is better because the subject is choosing to focus on segments that s/he can understand better in the first place. If this is true, Experiment 3 will show that a number of recognised items related to the most often repeated segments are present from the very first hearing of these segments.

(d) More than one of the causes listed in these three hypotheses could operate together, in a particular combination determined by the nature of the input, the task, and subject proficiency.

Table 9-14 below summarises the results of this analysis.

- *Recording D (burglary):*

Given that all move logs for recording D displayed very clear top-heavy patterns, none of them had any high-RR clusters located in the second half of the recording that could be analysed to test this hypothesis.

- *Recording C1 (interview):*

In protocols #7, #9 and #10, new decoding as well as inferencing processes were recorded in successive moves.

Protocol #6 was the only one where no inferences were found. Here, repeated playback was only used so that one of the subjects could pick up an item that had already been decoded by his partner:

Move 3	*Student A:*	– He's from Madrid
Move 6	*Student B:*	– What was the question?
Move 7	*Student A:*	– She said "de dónde eres"
Move 8	*Student B:*	– ¿Dónde es?
Move 9	*Student A:*	– De dónde eres.

(#6/3-9)

Overall, there was evidence to support both hypotheses (a) and (b) in recording C1.

As for hypothesis (c), there was no evidence of immediate item-recognition, but two pairs made remarks suggesting that it is pointless to repeat segments from which you cannot recognise any items: "Shall we try the bit at the start again? 'Coz I didn't understand much of that [=the end of the recording]" (#7/15); also found in protocol

Recording	Protocol	Most repeated segments	*Confirmed hypothesis* (a) More time for decoding	(b) More time for inferencing	(c) More items decoded in the 1st place
C1	#6	38-39	√ + Peer cooperation		
	#7	35-40	√	√	
		45-47	√	√	No word recognition > Not worth continuing
	#9	32-37	√	√	
		43-44	Actually focusing on adjacent segments (see ss.32-44 above) played back within the same moves.		No word recognition > Not worth continuing
	#10	36-38	√	√	Detects voice change > listen to question asked
B1	#11	48-55	√	√	√
		75-78	√		√
	#12	70-78	√	√	√
	#16	35-43	√		√
		55-66	√		√
F	#17	20	√		√
	#18	15-17	√		√
	#19	16-17	√		√

Table 9-14: Analysis of most repeated segments

#9: [tape stops and both laugh]: "Yeah I've lost it as well. - It's no good to me that" (#9/4), the task is terminated a few moves later: "Shall we keep going? - Well yes and no." (#9/9).

Finally, student #10 (working on his own) appeared to focus on a group of segments when he detected a voice switch that indicated a new question from the interviewer: "It's somebody else there... Where are you from? - Madrid" (#10/20). Relevant items

were decoded immediately (possibly as a result of focused attention triggered by the voice switch), which could also be interpreted as consistent with hypothesis (c).

- *Recording B1 (fat woman at doctor's):*

The three protocols will be presented separately as they represent three very different cases.

Protocol #11 contained supporting evidence for all three hypotheses - immediate word-recognition: "I don't know what she says about hair" (#11/1); inferencing: "It must be 'getting wet', I mean that would make sense" (#11/1), as well as new decoding as the segments are repeated: "sudar" (#11/2); "I don't like it [sudar[35]]" (#11/3). In fact, one new item is decoded each time the segments are played again. In the second cluster of segments (ss.75-78), repetition was used to check that an instantly decoded item ("She says ' I'll think about it' " - #11/3) had been correctly decoded: "Yeah: 'I'll think about it' "(#11/4).

Protocol #12 also supports all three hypotheses: immediate word-recognition: "cin-cuen-ta-mil" (#12/2); subsequent decoding processes: "sixty four million... or is it a thousand? Sixty four thousand pesetas" (#12/3); and inferencing: "That's not that much is it?... I don't really know that currency" (#12/3). Note that neither the L1 translation nor the attempted inference were accurate in this case.

Protocol #16 was produced by a highly proficient subject, talking aloud while working on his own. All the responses recorded were related to decoding processes, mostly represented by substantial "chunks" directly uttered in L1. The subject's strategies could be summarised into a three-stage process:

	Strategies/processes:	***Examples:***
First time heard:	• Gist decoded straight away, in the form of chunks directly uttered in L1.	"I was told if I eat I can... I've got to do exercise" (#16/1) "It's a revolutionary technique" (#16/1)
Second time:	• Checking original L2 wording (working "backwards")	"Técnica revolucionaria" (#16/4)
	• Refining previously uttered chunks	"I thought if I did exercise I could eat" (#16/3)
	• Decoding new isolated words (uttered in L2).	"las tablas" (#16/4)
	• Identifying gaps in detail information.	"cuánto vece" (sic.) (#16/3) [= how many times per week?]

[35] "Sudar": to sweat.

Third time:	• Filling in missing detail information.	[las tablas] > "tablets" (#16/5) "once a week minimum" (#16/5)
	• Strategies used the second time also continue here.	"Ah! It's a new technique from *Japan*!" (#16/5)
	• Already decoded items may or not be repeated this time.	

This case confirms hypotheses (a), since repetitions allow more chances for accurate decoding; and (c), since complete chunks were decoded directly into L1 during the first move in which the segments were played; but (b) is not confirmed here, since no overt inferencing was recorded.

- *Recording F (football match):*

All three cases examined (#17, #18 and #20) could be summarised into the same two-step strategy:

(a) Locate information relevant to the test questions by recognising key words.

(b) Decode the selected segment(s) by playing them repeatedly.

Protocol #17: "When is it taking place? did you get that?" (#17/1) > "I heard 'noches' at one point" (#17/5) > "Tomorrow evening" (#17/8).

Protocol #18: "he said something about 'el sol' ... It's very cold and sunny" (#18/1) > "Oh right, I think he says at the moment it's sunny but it might be cold for the match or something like that" (#18/4) > "Yeah it's gonna be cold for the match" (#18/5).

Protocol #20: "We definitely have some weather in there" (#20/1) > [they work on a different question for a while] "Can we go on to something else then? All right what's this?... The weather forecast." (#20/14) > "Oh hang on a second, there's something about 'sol' there" (#20/16) > "Good weather, 'algo frío' what's that?... Some cold I think" (#20/17) > [and the pair continue processing the same segments over the next few moves].

This two-step strategy means that hypotheses (c) and (a) apply one after the other in all the cases examined for recording F, but not hypothesis (b).

9.3.2.8.2 Discussion

The above results suggest that each of the three hypotheses applies under different conditions:

- *Hypothesis (a):* High levels of repetition always promote further decoding (evidence in all recordings).
- *Hypothesis (b):* High levels of repetition promote inferencing, but only if overt inferencing is needed (evidence in recording C1).
- *Hypothesis (c):* High levels of repetition are related to immediate recognition of discrete items when the subjects are listening for specific information (evidence in recording F).

One feature occurred in every single protocol that was considered for this analysis of the most repeated segments: some amount of observable decoding always took place when the segments were repeated.

This does not necessarily imply that further decoding processes will take place whenever a segment is repeated This analysis focused only on those segments that had been *most* repeated by each of the pairs. Furthermore, a tick in column (a) of the results table (Table 9-14, p.230) only indicates that in at least one of the repetitions at least one new decoding token was recorded. What the results imply is that, whenever a subject plays a section of the recording significantly more often than the rest, new decoding processes take place at some point between the first and the last time the section in question is played back.

Results for the two other features under examination (new inferences, and items recognised from the first playback) varied with the recordings used:

- *Recording C1 (interview)*

This recording produced instances of inferencing strategies in three out of four protocols. The only exception (#6) was considered to be an isolated case of repetition produced for a specific purpose in which inferencing was not required (helping a partner to hear a specific item). Since all the other protocols contained evidence of new inferences being made during repetition, it can be concluded that high levels of repetition in this recording co-occurred with the generation of new inferences. Given that the key to the topic of the interview was the speaker's occupation, it is not

surprising that the inferences made here mostly reflected attempts to activate any schema that would provide clues to fill this information slot (e.g. "I think that's what television was about. He works on the television" - #7/10).

As for items recognised from the very first hearing, no evidence was found with this recording. On two occasions (#7/15 and #9/3-4) the subjects even stated that they got nothing or very little out of the segments in question. It must be pointed out however that the present analysis does not consider the most repeated segments located in the first half of the recording. Hypothesis (c) might well be an explanation for the high-RR clusters found around segments 6-11 in this recording (see discussion p.209).

- *Recording B1 (fat woman at doctor's)*

This recording produced further inferences in only two cases, and the nature of these was quite different from the type of inferences found in recording C1. Upon closer examination, none of the inferences found here appears to be of key relevance to the global meaning of the recording ("Maybe ['sudar' means] hairdryers"- #11/2; "64 000 ptas, that's not that much is it? [...] I don't really know that currency."- #12/3)[36].

Unlike recording C1, recording B1 produced evidence of decoded items in the first playback for each of the five protocols examined. However, it would be hasty to infer that this phenomenon is related to playback frequency. Automatic decoding of entire chunks from the very first move has already been identified as a typical response throughout recording B1, and was found even in segments that had not been repeated at all.

In other words: the only feature that seemed directly related to repetition in recording B1 was the occurrence of further decoding processes.

- *Recording F (football match)*

The presence of recognised items in the first playback for all the cases using recording F could be explained in terms of a strategy typical of listening tasks that are cued by specific questions. Such a strategy has been referred to as "Step two" (locating relevant information) in section 9.3.2.4.2 above. The subjects search for

[36] According to the classification proposed by Trabasso and Magliano (1981), these two examples would be regarded as "paraphrase" and "association" respectively. Neither type of inferencing serves an integrative function (see p.42).

salient words related to the questions that they need to answer, and then repeat the relevant sections until all the information required is decoded. Since the general schema ("forthcoming football match") is already obvious and the information slots to be filled in ("who?", "when?" etc) are already supplied in the form of questions, no inferencing is required here.

9.3.2.9 Social-affective strategies

Hypothesis 9 proposes that social-affective strategies overtly affect the playback strategies that subjects use. The protocols were searched for any explicit evidence of social-affective strategy use. The main finding was that this data collection method elicited a large amount of valuable information on spontaneous social-affective behaviours. The students soon forgot that they were being recorded, and even when they did not ("Big Brother is listening to you" #9/7) , the physical absence of a researcher or teacher encouraged them to express their feelings more freely than in a conventional data-collecting interview. Negative reactions in particular would have been unlikely to come up in conventional introspective methods (e.g. remarks referring to the researcher who set such a difficult task: "Bloody woman!" #2/3, "Can we kill her?" #9/7). The language used was also free from register adjustments and allowed a good insight into the pairs' relationships: "—I don't know what he's saying... —Oh, don't be wet, come on!" (#2/6).

Despite temptations to report the vast array of social-affective behaviours found in the data, only those that are somehow related to PL strategies will be reported in this section.

9.3.2.9.1 Pair dynamics in task planning

Task planning is a metacognitive strategy. The way in which different pairs went about it however was largely related to the type of partnership that was established between the two participants.

In two instances (#4 and #13), the playback procedure was negotiated even before the first full run, in move "zero":

A: What do you want to do?
B: Just listen to it straight through.
A: Play it through all the way through first of all?
B: Okay.[Teacher can be heard saying "no notes".]
B: No notes? Okay. (#4/0)

More frequently, the pair would do this after the first full run: "Shall we listen in sections?" (#15/1); "—Play the first bit. —Right. [...] I'll give it to you to stop it. (#4/1). In most cases negotiation took the form of a simple suggestion to play back, readily accepted by the partner: "—Right let's play it again. —Okay." (#12/2). Suggestions to rewind were found throughout the task in every protocol.

In two protocols (#14 and #21) one of the subjects was found to be making all the decisions on playback procedure:"OK, let's listen from the beginning" (#14/5); "Yeah it's 'ochenta mil' isn't it, it's not 'ocho mil'. Just rewind it and check it's not 'ocho mil' " (#21/4). These cases are examined in section 9.3.2.9.3 below.

In a few instances, one of the partners took action to ensure that the other partner's attention was not distracted from the goal in hand: "—'celebrando'... —Yes we got that bit. [Let's now work on] 'De repente' " (#2/9); "Just fill in the last bit" (#21/9); "Let's just carry on with it and we'll sort that back" [sic] (#5/2).

9.3.2.9.2 Reactions to input exposure

The first contact with the input triggered immediate reactions to speech rate and auditory conditions ("What a relief it stopped!" #4/1), which have already been reported in section 9.3.2.1.1 (p.194).

Evaluative comments on a subject's own performance were also a commonly found as an immediate reaction. They could either be positive: "I understood the first few sentences" (#9/1), "well I got the first one!" (#21/1); or more often negative "I don't really quite catch a lot to be honest" (#1/1- see also #2/1, #5/1, #9/2). Although self-evaluation itself is a metacognitive strategy, self-evaluative responses tended to be marked by an affective component ("I'm really getting nothing out of this"- #5/21), which is why they are examined in this section.

Immediate reactions were found in one form or another for challenging recordings such as D (burglary) and C1 (interview) - with only two exceptions: #6 and #8. No such reactions were found for easier recordings such as B1 (fat woman at doctor's) and F (football match). Wherever they occurred, they were always found before any cognitive processing of the input had begun. The only exceptions to this were two cases (#2 and #5) in which a few L2 words were repeated on-line while the tape was still playing. Even in such cases, the reactions were voiced immediately after stopping the tape, and only then did off-line processing begin.

Reactions to speech rate and listening conditions were most frequent at the very beginning of the task (after the first full run, or even the first short move), but they could also be found at other points of the recordings (#7/4, #7/12, #17/3, #21/13) and particularly at the end of a run (#2/2, #5/21, #6/2, #7/2, #7/15). Although two protocols did show such reactions later in the task for recording F, there were still no cases found for recording B1.

Self-evaluative comments between two runs were found with all recordings: recording D (#2/2, #2/18), recording C1 (#5/5, #5/21, #6/2, #6/14, #7/2, #7/15, #9/7, #9/9), recording B1 (#11/4, #13/2, #14/5, #15/1), and recording F (#17/8).

Looking at the overall affective response (including other behaviours that are not reported here such as laughter, deliberate digressions, etc.), recording F (football match) produced few affective responses in comparison to recordings D and C1. Recording B1 (fat woman at doctor's) produced even fewer, and the only affective behaviour that was observed recurrently for this recording was occasional laughter while the tape was playing, in response to amusing stereotypical reactions from the speakers (#11/3, #11/3, #12/2, #15/2, #16/1, #16/4).

9.3.2.9.3 Types of partnership between task participants

The type of relationship that was established between task partners was found to have visible effects on the playback strategies that they used in a number of instances. Three possible types of partnership were identified:

Cooperative and balanced

Pair #8 was the clearest example of this type of partnership. The pair comprised two female students, one English, one Greek. The Greek student found it easy to recognise words in L2, whereas the English speaker experienced segmenting difficulties "I can't tell where one word finishes and the next word starts" (#8/42). Yet her attitude to her partner's advantage was generally positive "Wey! I didn't understand that" (#8/6), "Did he? Wow!" (#8/30). The Greek student would simply locate and play back the relevant segments for her partner until she either heard the words in question (#8/6) or took her partner's word for it (#8/24 below). Once the decoded L2 words became available to both, the English student actively contributed to further processing by providing L1 equivalents, grouping given information and making inferences:

Greek student: — Arquitectura. You know?

English student: — Oh, right? I missed that! He's an architect in the morning, TV presenter in the afternoon.

(#8/24)

This strategy produced a series of many short moves, resulting in a very slow overall progression.

Similar cooperative partnerships were found in all the other cases that had been identified as "slow overall progressions" in section 9.3.2.5.2 (p.219); namely pairs #2 and #20 (the remaining slow progression was from #10, but it is not relevant here since the student was working on his own):

Pair #2 shows evidence of mutual support: "—I don't know, I don't know. —Don't worry" (#2/7); "—We haven't got a clue [...] —We *have* got a clue." (#2/19). A "keep trying" approach to difficulty can also be found throughout the task: "Well, we haven't got a clue anyway, so we may as well listen to it anyway" (#2/19); "Go on. Play, play!" (#2/15). Mutual support is equally clear in the case of pair #20 after a problematic passage:" —What was that? — [laughs] I don't know! — O.K, we're doing well here. It's impressive." (#20/20).

Cooperative and un-balanced

Like the cases discussed above, pair #21 used a cooperative approach to the task, but here one of the students (student A) was more proficient than the other (student B). As a result, most decisions on playback procedure were made by student A. "Shall we just listen to it again, just to get that? Rewind the lot and just get it all [...] and then just double-check as you listen to it" (#21/9). The suggestion almost has instructional undertones, and this "teaching" role is even more evident when B is writing down the answer to question 5, and A spells out an abbreviation for him:

A: — P-T-S, just put P-T-S... "pesetas".

B: — E-T-S ?

A: — P-T-S, just P-T-S.

B: — Oh, P-T-S !

A: — It's just like... "pound" you know.

B: — Oh right! [...] ... O.K.

(#21/5)

As was the case for pair #8, the more proficient student offers to repeat discrete sections of the input so that his partner can check for himself the information that A has just provided: "Perhaps you wanna listen to it again" (#21/8). He even takes care to stress the difficulty of the passage so that B is not discouraged "He talks very quickly doesn't he [laughs]"(#21/8).

Unlike the previous examples, this pair's move log did not display a pattern of slow overall progression.

Non-cooperative

In protocols #9 and #14, there appeared to be a mismatch between the partners' level of commitment to the task.

In pair #9, one of the partners (student B) expressed negative evaluations of his performance after almost every move: "I just lost it there" (#9/2), "I totally lost gist of it" (#9/3), "It's no good to me that" (#9/4). His partner (student A) went along with such comments occasionally: "Yeah I lost it as well" (#9/4), but he tried nonetheless to process some of the information in the recording:

A: — Is he not going on about something about local TV and presenting and stuff?

B: — Yeah yeah. But like, I could listen to that about a hundred times, and I'm not gonna be able to [make head or tail] of it.

(#9/5)

Instead of the "keep trying" approach that was found in cooperative partnerships, a "no point" approach was adopted here by student B early on in the task: the segments to which the remark cited above refers had only been played twice when the remark was made. The following conversation took place only 11 segments into the recording (which was 66 segments long altogether):

B: — You're all right with that bit?
A: — [laughing] It's a bit fast isn't it!
B: — I just lost it there.
A: — Yeah.
B: — That's what happened last time, on the whole one [=a previous class test].
A: — Hum.
B: — I listened to the first three sentences, and then the rest of it was just like
A: — And then they're up to speed by the fifth sentence [laughs]

(#9/2)

Student B did suggest repeated playback a couple of times (#9/6, #9/7), but by move 7 his attitude was overtly hostile to the task, asking to "kill" the researcher ("Big Brother") and mocking the speaker's voice (#9/7).

Student A tried to proceed with the task, but failed to get B's support:

A: - Er... I'm not quite sure why he wants to get to Madrid...
B: - I don't think we'll ever know. (#9/7)

Eventually, his final attempt was cut short:

A: - Shall we keep going?
B: - Well, yeah and no. [silence] Any comments there? Anything about the tape?
A: - Apart from the fact that I didn't understand it all, yeah.
B: - Yeah! [starts humming] (#9/9)

The last remark made by A in this example shows that negative self-evaluation was the final outcome for both partners.

The other case of non-cooperative partnership was pair #14, comprising two non-English native speakers: a Greek female (A) and a Turkish male (B). The entire playback procedure was decided by B, who was clearly the most assertive:

B: - Much better this time, isn't it? You go first.
A: - No!
B: - OK [starts reporting] (#14/5)

As was the case for pair #9, A's efforts to continue processing the input are cut short, as B makes a unilateral decision to terminate the task:

A: - Er... puede.
B: - It could be puede, but I didn't get it. Er... that's it. Shall we just...
A: - Frío.
B: - Maybe, maybe. Shall we just fill the paper? OK. We're done. (#14/5)

Unlike with pair #9, no negative self-evaluations were found here, and the evaluations made by B showed constructive attempts to analyse the difficulties experienced: "Well I remember, but... just the words, not exact... opinions. I don't know, we seem to remember more about that part I think, the beginning part. Er... so. It's all I can remember" (#14/4, note that the use of "we" is taken for granted). The evaluation that followed the second run was even positive: "Much better this time isn't it?" (#14/5).

The cause for early termination in the case of pair #14 could be a low target level of comprehension on the part of the most assertive student. After the first run, he simply

reported: "OK. The thing is like she's a doctor, she goes to the doctor, she's got a headache. Am I right? Em.. 'muchos muchos' headache. [laughs] I'm done." (#14/3). The "headache" inference was never checked, and the statement "I'm done" at such an early stage in the task turned out to be more than a simple joke. The attitude here could be summarised as "no need", rather than the "no point" attitude adopted by pair #9.

The move logs in both non-cooperative cases displayed fast overall progressions. This is not particularly significant for #14, since all progressions were fast for recording B1. However, move log #9 was identified in section 9.3.2.4.2 as the only case in which a fast overall progression had been used through both the first and second halves of recording C1.

9.3.2.9.4 Discussion of social-affective strategies

The fact that students worked in pairs promoted the use of social strategies, which would not have been possible if they had been working individually. This procedure also provided them with a genuine interlocutor with whom any affective reactions to the task could be shared in an equal-to-equal situation.

In a recent study, de Guerrero and Villamil (2000) conducted an analysis of the moment-to-moment changes in task partners' behaviour while they worked collaboratively on a writing task. The task in question consisted of revising jointly a written text produced by one of them in L2. The researchers focused on the interactions of one single pair of students (2 male intermediate ESL college learners, native speakers of Spanish), who were selected because "their interaction was sufficiently rich and varied to allow the observation of a wide range of behaviours..." (de Guerrero and Villamil 2000:55). The authors argue that in this type of set-up, both learners operate within their respective zones of proximal development. The process of "mutual scaffolding" in which they engage is similar in many ways to the scaffolding that takes place when an adult helps a child perform a task within the child's ZPD (Vygotsky 1978). Many of the cooperative behaviours that the researchers observed during the revision task were very similar to those found in the present study during the joint listening task performed in Experiment 3. Among these were:

- the use of "we" throughout the task, and especially at moments of high activity, which indicates that the participants are working together on the task as a single unit with a shared "regard".
- the use of humour as a means of relieving stress, but also to consolidate a state of mutual cognition called "intersubjectivity", that is essential when working within the ZDP. Since it is not directly related to playback strategies, this strategy is not analysed in detail here, but it was found in many protocols and is believed to play an important role in facilitating bonding between task partners.

Pair dynamics in task planning

De Guerrero and Villamil identified an initial stage of metacognitive planning, in which task participants negotiated a procedure to follow: "First I am going to read your composition" (de Guerrero and Villamil 2000:56). This strategy is very similar to the planning strategies observed in the present study, in which partners agreed on a playback procedure either before commencing the task or immediately after the first run ("Shall we listen to it in sections?" - #15/1). Exposure to the input is the necessary first step to processing any message. Therefore it is not surprising that metacognitive strategies related to the manner in which such exposure ought to take place should spontaneously occur at the beginning of the task.

Reactions to input exposure

One finding in Experiment 3 was that there were fewer overt affective reactions when the task was not very difficult (recordings B1 and F). Most affective responses were found in recordings C1 and D, often in the form of negative evaluations of one's own performance. Acknowledging difficulty and seeking the partner's agreement on the "unreasonable" demands of a task could be a face-saving strategy aimed at justifying what a subject may regard as unsatisfactory performance. Justifying poor performance is a strategy commonly used in any area of human activity (e.g. a person trips on a trailing cable and immediately argues: "I didn't see it!"). When listing social-affective strategies, the literature tends to focus on strategies leading to success, such as "lowering your anxiety" (O'Malley and Chamot 1990; Oxford 1990), "cooperating with others" and "encouraging yourself" (Oxford 1990). Little attention is given to less constructive strategies such as face-saving acknowledgements of failure and refusal to cooperate (the "no point" approach is a clear instance of an affectively

motivated avoidance strategy). The findings of this study show that such strategies are sufficiently widespread to deserve greater attention. Acknowledging and understanding "negative" strategies should also enable us to develop teaching methods that minimise their adverse effects on performance and motivation.

The fact that negative judgements were found more often than positive ones does not necessarily indicate that the subjects had a negative perception of their performance overall. It is likely that frustration leads to more verbalised responses than satisfaction. When processing successfully, the subjects may simply get on with the task in hand, either too busy or feeling no need to voice their emotions. This is confirmed by the lack of affective responses found for recordings B1 and F. Besides, it has been argued that there is no such thing as "100% comprehension" (especially in L2), only "reasonable interpretations" of the input (Brown and Yule 1983b:57). A "good language learner" is said to tolerate reasonable levels of uncertainty and still be able to experience a general sense of success (Rubin 1975). This co-existence of partial failure with an overall sense of success was found in some cases ("We've got it all. I reckon. Apart from 'sudar'" - #11/4). In their study, de Guerrero and Villamil observed a tendency to over-simplify the final evaluation of the task after completion: "the only thing I did wrong was the (...) spelling mistakes" (de Guerrero and Villamil 2000:63). Such tendency to minimise the difficulties experienced was interpreted by the authors as a "praise and encouragement" strategy.

In terms of playback strategies, the social-affective strategies found in Experiment 3 can be summarised as follows:

- When the input presents some difficulty to the listener, affective reactions are voiced during the first few moves or in the first run, before controlled processing begins.
- Evaluations of one's performance often occur at the end of a run. This could be because until the run is completed, subjects are generally busy processing the message. Their attention is focused on cognitive processing and cannot afford to be diverted onto other areas.

Types of partnership between task participants

De Guerrero and Villamil (2000) examined the existing relationships between students working in pairs on a writing task. Unlike Experiment 3 in this study (in which

learners paired together were of similar proficiency levels and had identical roles at the outset), the set-up that they used allocated the role of "reader" to the more proficient subject in the pair. Even so, they found that the less experienced "writer" gradually assumed more responsibility in pointing out problem areas and proposing solutions.

The present study offered a better opportunity to observe a range of role-related behaviours, since the roles were not pre-determined and a relatively large sample was used (de Guerrero and Villamil only studied one pair). The different types of pair dynamics identified in this experiment have been labelled "cooperative and balanced", "cooperative and unbalanced" and "non-cooperative".

The relationship between task partners was found to influence playback strategies in the following ways:

- When the relationship between task partners is cooperative and balanced, the pace of progression is slow; whereas non-cooperative partnerships result in fast progressions.
- One observed cooperative strategy consists of one subject playing back selected segments that s/he has managed to decode, so that his/her partner can locate them and decode them as well (#6/6-9; #8/6; #21/8).
- Playback procedure is generally negotiated between task partners, though in some cases the more assertive partner takes the initiative. When playback decisions are made by a proficient listener (#21), comprehension may be facilitated. When they are made by a non-proficient listener (#14), lack of balance in decision-making has negative effects on performance.

Some of the features observed by De Guerrero and Villamil (*ibid.*) were also found in this study. The most relevant are listed below:

- The authors noted an implicit role-allocation as "expert" and "novice" through the explicit roles of "reader" (who proposes the revisions) and "writer" (who may accept, reject, or elaborate on the proposed changes). The attitude of the "tutee" in this dyad was overall one of openness to the changes suggested by his/her partner. This was also the case in the "cooperative and unbalanced" partnership found in the present study (pair #21).
- An ability to switch between the roles of "expert" and "tutee" on the part of both subjects (even in the pre-defined set-up that was used in the writing study),

especially as the task progressed and partners became more "in tune" with each other. For instance, the "writer" was found to spell out an unknown word for the benefit of the "reader" (de Guerrero and Villamil 2000, "episode14"). These role switches were a common feature of "cooperative and balanced" partnerships in Experiment 3 (e.g. pair #8).

- A responsive attitude (termed "contingent responsivity") to the behaviour of the "tutee" on the part of the expert, being attentive to any cues that would indicate the current affective or cognitive state experienced by the less experienced partner. This prevents the scaffold from "collapsing" as a result of frustration. There were several instances of this attitude in the present study (e.g. #20/20; #21/8).
- Occasional episodes of "regression", where the more assertive subject in the dyad persuades the other to adopt an incorrect solution. This only occurred rarely in de Guerrero and Villamil's example, but a few instances were observed in the present study (see the discussion of pair #14 below).

In terms of playback strategy, it would seem that

- balanced cooperative partnerships (#2; #8; #20) result in slower progressions (due to a "keep trying" attitude reflected in effective scaffolding strategies);
- non-cooperative partnerships (#9; #14) result in faster progressions and early termination of the task (due to a "no point" or "no need" attitude, in which the scaffold collapses: the underlying social-affective strategies are either not used, or aimed at a purpose other than learning).

The attitudes found in this case were a result of pair dynamics, but it could be hypothesised that, to a certain extent, similar types of attitude might affect the playback procedure of students when they work individually.

The case of pair #14 needs to be considered separately, since both partners were non-native speakers of English: a Greek female (student A) and a Turkish male (student B). Although the recording did not present difficulty to any of the other pairs, this pair failed to reach an acceptable meaning hypothesis. Student B took the leading role right from the beginning of the task, and appeared satisfied with his first interpretation. Student A tried to contribute a number of decoded items, but none of these were utilised by B, who did not allow A any opportunities for further elaboration. This pair's failure to complete the task successfully appears to be due to insufficient processing rather than difficulties with processing itself. Such an approach was

deliberately promoted by B, and remained unchallenged, since student A lacked the assertiveness required to pursue further processing without her partner's support. The lack of cooperation in pair #14 may also be due to the fact that both partners were operating in English, a foreign language to both of them. Besides the cognitive difficulties related to lack of proficiency in English (increased processing load, lack of English equivalents for Spanish words in the mental lexicon, etc.) , the affective side of communication may also have suffered from the use of a language that was foreign to both partners. De Guerrero and Villamil (2000) argue that the use of L1 as the chosen language for interaction during the task constitutes "a critical psychological tool that is essential for collaboration" (de Guerrero and Villamil 2000:64). A further variable in the form of historical/cultural factors is also very likely to have contributed to the lack of cooperation. In retrospect, a Greek female and a Turkish male probably constituted an unfortunate match. Nevertheless, the pair's behaviour provided a useful insight into the strategies used by non-cooperative task partners. The relevance of "negative" strategies will be discussed in the next chapter.

9.4 Summary of Experiment 3 findings

This section summarises the specific conclusions derived from the above discussion of protocol analysis results. It also relates these findings to global patterns previously found in the comparative analysis of move logs from the four recordings (section 9.3.1, p.186 ff.). The general implications of the specific conclusions listed below will be examined in section 10.3 of chapter 10.

9.4.1 The first run in playback listening tasks

- When input is "difficult", *reactions to speaker's accent and delivery* occur after the first run, but there are no such initial reactions when input is "easy"
- *Overt schema-activation strategies* occur in the first run when subjects cannot activate a suitable schema from their first exposure to the input. The *pace* of the first exposure does not appear to be relevant.
- The strategy of *locating information that sounds relevant* is used at the beginning of a task, *regardless* of the pace of progression used in the first run, and provided that the following conditions are both met:

    1. a suitable schema has been identified

2. the specific information required in order to fill in the slots in that schema has not yet been decoded

9.4.2 The last run in playback listening tasks

- *Hypothesis-checking* is a typical strategy in the final run of a listening task, regardless of the pace of the final run in question.

9.4.3 Top-heavy patterns

When they start working through a recording in L2, *listeners first need to achieve a minimum level of schema activation* in order to account for one (and ideally more) of the following aspects:

- the situational schema of the message (e.g. "at the doctor's")
- the formal schema followed in the message (e.g. "problem-solution")
- the topic of the message (e.g. "being overweight")

From that point (but only then) they can start using any clues available to help fill in the information slots assumed to be contained in the active schemata.

- If for any reason the necessary schemata are activated straight away, they can proceed to search for relevant information throughout the recording to fill in as many slots as possible.The resulting playback strategy is *not top-heavy*, as shown in the case of recordings B1 and F.
- On the other hand, when subjects experience difficulty finding suitable schemata, they try to decode as much information as possible from the beginning of the recording by means of controlled bottom-up processing. This produces the type of *top-heavy* playback patterns that were consistently found for recording D.

The need to rewind right to the beginning before the end of a run is more likely to arise when the flow of processing is significantly delayed by a problematic segment. If progression through the task is smooth, there is no need to stop in the middle of a run and start afresh from the beginning. This accounts for the *gradual decrease in playback frequency throughout more difficult recordings.*

9.4.4 Recording F (response cued by questions):

The initial steps seemed to be common to all the cases examined:

Step 1: read the questions

Step 2: listen to the recording straight through, with attention to any information relevant to the questions

The outcome of this step is twofold:

- Some information is processed straight away, ready to be utilised in response to the relevant question.
- Some of the output is in the form of partly processed information, which is used to locate the part of the recording in which the answer to a particular question can be found.

Step 3: utilise the information that is fully processed (i.e. answer the relevant question from the answer sheet).

Step 4: fill in the gaps.

It is at this stage that the strategies found begin to differ across the sample: subjects may follow:

- the sequence in which information is presented in the input
- the sequence in which information is prompted in the questions
- a sequence determined by incidental recognition of relevant items in the input

Saliency and *complexity of information* were also found to be determining factors for the playback strategies used.

9.4.5 Fast overall progression

The processing strategies underlying fast overall progressions were found to be closely dependent on the nature of the input. Three different scenarios for fast overall progressions were identified:

1. Subjects *process the input automatically* when input is "easy"
2. Subjects *cannot process the input* when input is "difficult"
3. Subjects *will not process the input*, whether input is perceived as "easy" (the "no need" approach) or "difficult" (the "no point" approach)

9.4.6 Slow overall progression

High numbers of decoding tokens were found in slow overall progressions, which indicates that *a larger amount of controlled decoding processing* was taking place. Hence the need to work intensively on a small number of segments at a time, since controlled decoding proceeds more slowly than automatic decoding. However, the findings were regarded as "encouraging" rather than conclusive, given the small number of cases of slow overall progressions available in the sample.

9.4.7 The a-b-a strategy (fast-slow-fast)

Whenever the a-b-a strategy is used, the processes and strategies found in each of the three stages are *largely the same as those proposed in the hypothesis, but they may be affected by input-specific and task-specific factors*, such as the availability of global schemata, the amount of input that can be decoded automatically, the availability of clues in the form of questions, and possibly a number of other factors not yet identified.

The original hypothesis consisted of:

Stage 1 (Fast run at start):

- *Getting acquainted with the speaker's accent and delivery*
- *activating relevant schemata*
- *locating information that "sounds" important*

Stage 2 (slow run in the middle)

- *Bottom-up processing more intensive*

Stage 3 (fast run at end)

- *Hypothesis-checking*

In the single occurrence of a-b-a pattern found in this study, the actual playback pattern obtained differed slightly with the one proposed in the hypothesis. Minor differences were observed in the form of:

- No evidence for the "locating relevant information" strategy in stage 1.
- During stage 2, intensive bottom-up processing gave way to global inferencing (a strategy not envisaged in the hypothesis), which was then applied to any items subsequently decoded throughout this stage.

- Stage 3 was used not only for checking previous hypotheses, but also to group them all into a final interpretation for the entire recording.

In conclusion, the strategy is not a typical behaviour (only one instance found), but there is ground to assume that it could prove to be a highly effective playback strategy, provided that it is used in combination with effective processing strategies that are probably task-specific (hence the variations described above).

9.4.8 Playback frequency and recall

Each of the three hypotheses applies under different conditions:

- *Hypothesis (a):* High levels of repetition always promote further decoding (evidence in all recordings).
- *Hypothesis (b):* High levels of repetition promote inferencing, but only if overt inferencing is needed (evidence in recording C1).
- *Hypothesis (c):* High levels of repetition are related to immediate recognition of discrete items when the subjects are listening for specific information (evidence in recording F).

9.4.9 Social-affective strategies

Playback procedure is generally negotiated between task partners, though in some cases the more assertive partner takes the initiative. When playback decisions are made by a proficient listener, comprehension may be facilitated. When they are made by a non-proficient listener, lack of balance in decision-making has negative effects on performance.

When present, negotiation of overall playback strategy tends to occur at the beginning of the task. Exposure to the input is the necessary first step to processing any message. Therefore it is not surprising that metacognitive strategies related to the manner in which such exposure ought to take place should spontaneously occur at the beginning of the task.

In terms of playback strategies, the social-affective strategies found in Experiment 3 can be summarised as follows:

- When the input presents some difficulty to the listener, *affective reactions are voiced* during the first few moves or in the first run, *before controlled processing begins.*
- *Evaluations of one's performance often occur at the end of a run.* Until the run is completed, subjects are busy processing the message: their attention is focused on cognitive processing and cannot afford to be diverted onto other areas.
- *Negative evaluations* are more frequently voiced than positive evaluations.

The relationship between task partners was found to influence playback strategies in the following ways:

- When the relationship between task partners is cooperative and balanced, the pace of progression is slow; whereas non-cooperative partnerships result in fast progressions.
- One observed cooperative strategy consists of one subject playing back selected segments that s/he has managed to decode, so that his/her partner can locate and decode them as well.

9.4.10 Global patterns observed in the four recordings

At this point, it should become possible to resume the discussion of the differences reported in section 9.3.1.3 (p.191) between global playback patterns in the four recordings used.

In very broad terms, it appears that the relative difficulty of a given recording determines which playback strategies subjects choose for tackling it. However, a simple glance at Table 9-15 shows just how many inter-connected factors need to be considered when trying to establish the level of difficulty that a given recording presents to a specific population. Many other factors could have been added (see p.50) but in this case only those most relevant to the discussion of playback listening strategies have been included. The table is divided into three main sections:

- *The top rows (A and B):*

Rows A and B summarise basic input-specific features such as the nature of the information presented ("static", "dynamic", "abstract")[37], speech rate and discourse modality (scripted/uscripted). The product of these input variables considered in combination appears to confirm the preliminary grading of the recordings proposed in section 9.3.1.3. In other words: B1 should indeed be the easiest, F should present a moderate level of difficulty, whereas D and C1 should be the most difficult recordings.

- *The middle rows (C to F):*

An important factor appeared to be whether or not listeners were able to establish from the start a valid schema upon which to base further processing of the input. This aspect of the task is examined in the middle section of Table 9-15.

- *The bottom rows (G and H):*

Row G summarises the distinctive global patterns that were identified earlier in this chapter (Table 9-2), and needed to be accounted for by means of protocol analysis. The bottom row (H) summarises possible explanations that emerged from the discussion of protocol analysis results.

[37] See Brown and Yule's (1983b) categories (p.50).

	Recording D *Burglary at speaker's house*	***Recording C1*** *Interviewing a local TV presenter*	***Recording B1*** *Fat woman at doctor's*	***Recording F*** *Forthcoming football match*
A) ***Nature of input:*** See Brown and Yule's (1983b) categories (p.50).	*Dynamic* (narration)	• Part 1: *Static* (description) • Part 2: *Abstract*	*Static* (description /instruction)	*Static* (description)
B) ***Other features of speech affecting difficulty:***	• Speech rate: moderate/fast • Unscripted monologue	• Speech rate: fast • Unscripted interview	• Speech rate: moderate. • Unscripted role-play by two language teachers	• Speech rate: fast • Scripted report from off-air radio
C) ***Schema available prior to listening:***	None	None	None	*"Forthcoming football match"* + related info
D) ***Earliest clues to become available for schema activation:*** (as found in protocol analysis results)	A few unconnected words from ss 1-6: *"friends", "house", "celebrating", "medic", "end of degree".*	First few turns in the recording (ss.1-12): *"My name is...", "I live...", "What is your occupation?"* etc.	The opening utterance: *"Doctor, I have a problem"* (s.1), followed by *"I eat too much"* (s.3).	Five comprehension questions given prior to listening.
E) ***Schema readily available from segments decoded early on:*** (as found in protocol analysis results)	None Even the schema *"graduation party"* has to be reached through inferencing, given the subjects' proficiency level.	• Format: Interview • Initial topic: speaker's personal details	• Situation: doctor's surgery • Topic: overeating	No decoding required for initial schema activation: the questions already tell us what factual details will be mentioned at some point in the recording.
F) ***Relevance of initial schema to the global meaning structure:***	0% relevant *"Graduation party"* is only of peripheral relevance to the main *"burglary"* schema.	50% relevant *"Interviewee giving personal details"* is only relevant to the 1st half of the recording	100% relevant *"Seeing doctor about overeating problem"* is directly relevant to the global meaning of the entire recording	100% relevant The 5 questions are directly relevant to global meaning and to specific information through the entire recording.
G) ***Global playback patterns observed:*** (from move log data)	**1.** Avge. move is a few segments long (i.e. pace is slower) **2.** All subjects used distinctly top-heavy patterns **3.** All subjects started with a fast run	**1.** Avge. move is a few segments long (i.e. pace is slower) **2.** Playback frequency gradually decreasing	**1.** Avge. move is many segments long (i.e. pace is faster) **2.** No gradual decrease in playback frequency	**1.** Avge. move is a few segments long (i.e. pace is slower) **2.** Playback frequency moderately decreasing **3.** All subjects started with a fast run
H) ***Related listening strategies:*** (from discussion of protocol analysis results) The numbering used in this row refers to the numbering of playback patterns in row G.	**1.** Processing is largely controlled **2.** Global schema not obvious: looking for it in the opening segments **3.** Scanning for a global schema first of all	**1.** Processing is largely controlled **2.** Two reasons possible: – Increasing difficulty causes regular returns to the start. – Part 2 too difficult: no point attempting to process it.	**1.** Processing is largely automatic **2.** Not difficult means no need for regular returns to the start	**1.** Processing is largely controlled **2.** Moderate difficulty: occasional returns to the start in order to find previously located information slots **3.** Locating any info. relevant to the questions given

Table 9-15: A comparison of the four recordings used

The key formula determining global playback strategies appears to include the following factors:

1. The nature of the clues available for establishing an initial schema (see rows C and D on Table 9-15).
2. The amount of decoding/inferencing that subjects need to perform in order to establish such schema (row E).
3. The relevance of the initial schema to the global meaning of the recording (row F).
4. The amount of information that subjects can obtain by means of automatic processing alone (Typically: most of recording B, little/none of C1 and D, a few items in F).
5. The amount of information that subjects can obtain by means of controlled processing (Typically: some parts in the first half of recording C1; most of recording F; and a minimal fraction of recording D, mostly obtained through inferencing).

Obviously, none of these factors can be defined solely on the basis of input features. Other dimensions also play a crucial role. For example, task features (such as the presence/absence of comprehension questions prior to listening) may determine how early in the task listeners are able to activate a valid schema; individual variables (such as proficiency) will determine the amount of information that a given subject will be able to obtain through automatic processing. These two examples illustrate the complexity of interaction between the many variables affecting playback strategy.

The three groups of phenomena reported in rows G and H of the table can be justified as follows:

1. Differences in the pace of progression (much faster in B1 than in the others):

Moves were much longer (indicating a faster progression) in recording B1 because most processing was automatic, whereas in the other recordings more controlled processing was required. Even recording F required listeners to perform a certain amount of controlled processing in order to fully decode the specific information required. The availability of a comprehensive schema certainly helped (the fact that

all pairs completed the task is indicative), but fast speech rate and sometimes complex syntax prevented subjects from relying on automatic processing alone.

2. *Decrease in playback frequency as the recordings progress / top-heavy pattern in recording D:*

Recording D is an extreme case (the most difficult of all four), because the very few words that listeners were able to recognise from the first few segments were not related to the main schema underlying the global meaning of the recording. Even the marginally relevant schema "graduation party" could only be reached through controlled processing, involving a small amount of word recognition and a large amount of inferencing. With this recording, subjects hardly ever reached the next stage in which specific information is decoded and integrated into a global meaning hypothesis, and consequently they kept trying to make sense of the beginning of the recording. This explains the distinctive top-heavy patterns found in all move logs for this recording. Using such a strategy was ineffective in this instance, given the lack of relevance of the opening segments to the rest of the story.

Recordings C1 and F also showed negative correlations between segment number and repeat rate (RR), indicating that playback frequency was greater towards the beginning than towards the end of these recordings. Unlike recording D, where all the segments defined as "high RR" (see p.209) were located exclusively in the first half, these two recordings showed a more gradual decrease in RR. Such negative correlation was most pronounced in the most difficult recordings (D and C1), less in the moderately difficult (F), and non-existent in the easiest (B1).

When sequential progression through the recording is smooth (as when most processing is automatic), there is little need to rewind. On the contrary, if problems are encountered at some point, listeners will rewind back to the beginning as they try again to find a link between the problematic section and whatever they have understood so far. The more problems found, the greater the tendency for RR to decrease through the recording. Primacy could also cause subjects to focus the best of their efforts on the beginning of a recording, whilst fatigue would cause a gradual decay of their cognitive and affective involvement in the task towards the end. The more effort it takes to make progress, the greater the fatigue, and the more evident the effect will be. Protocol data for recording C1 also showed that some pairs

resorted to much faster progressions (i.e. fewer moves/ repetitions) in the second half because the final section was simply too difficult to process.

3. Use of a fast run at the beginning of the task by all subjects in recordings D and F:

The fast opening run that was observed in all move logs for recordings D and F occurs for different reasons in the two recordings:

In recording F, a fairly comprehensive schema is already in place even before the first exposure to the input, based on the questions provided prior to listening. Therefore the next step for listeners is to scan the recording in order to locate specific information and fill in the slots outlined in the schema.

In contrast, listeners have no schema whatsoever on which to base their interpretation of recording D. Their first scan through the recording will be in search of any clues to a workable schema hypothesis.

Both approaches had been predicted in hypotheses 1c (1st run used for locating information, as in recording F) and 1b (1st run used for schema activation, as in recording D) respectively.

This concludes the discussion of global playback patterns across the four recordings. The findings of this section will be synthesised and evaluated in the next chapter.

10. Conclusions

The original purpose of the study was to answer the following questions:

1. What are the specific features of playback listening (PL)?
2. How can PL be used for research on listening processes and strategies?
3. What does initial PL research tell us about L2 learners' listening processes and strategies?
4. How can PL be used in order to develop listening skills in L2?

This chapter examines the extent to which the present study has succeeded in finding answers to those original questions, and raises a number of further questions that have emerged in the process:

1. The first section will discuss the features of PL that were identified in chapter 3.
2. The second question will be addressed in the section entitled "Methodological advances from PL research", which presents an evaluation of the research instruments developed for this study.
3. The third section ("L2 listening processes and strategies in PL") addresses question three and evaluates the findings of this study with regard to specific playback strategies. Areas for further research are also discussed.
4. Finally, a number of recommendations for the teaching and testing of listening skills will be proposed in the final section, in response to question four.

10.1 Features of PL

- *What are the specific features of PL?*

The present study introduces the notion of playback control as a relevant feature in a listening task involving recorded input. The definition of *playback listening* proposed on p.54 clarifies this notion in relation to commonly related concepts such as "noninteractive", and "recorded". The two further definitions provided for *playback listening strategies* and *playback strategies* (p.55) integrate the feature "playback control" within the general framework of listening strategies. Whilst many PL strategies are shared with listening tasks that do not involve PL and are not

necessarily overt behaviours in themselves, this study has succeeded in identifying a limited number of overt behaviours that indicate the use of certain PL strategies. The third notion introduced was that of *playback strategies.* Unlike PL strategies, these always constitute overt behaviours that can be examined by an external observer.

A number of researchers have argued the need for empirical studies that use observational criteria in strategy research (Rees-Miller 1993; Seliger 1983). It is believed that the newly introduced concepts are a step in the right direction. PL-related notions should also assist in the analysis and discussion of information obtained from any experiments that use recorded input, whether they investigate strategy use or other aspects of listening.

The issue of authenticity was also discussed with regard to the status of PL in foreign language instruction. It was concluded that, although PL is not the most typical mode of listening in real life situations outside the instructional environment, it does often represent the real life experience of the learner more accurately than face-to-face interaction (see section 3.1.1.1). In that sense, PL can be regarded as an "authentic" activity. The same argument applies to the skills and strategies involved in PL. Some of these (e.g. explicit observation of language use, controlled decoding of short repeated segments) may not seem "authentic" in terms of the skills that would be used outside the classroom. They are nonetheless pertinent, and potentially useful skills and strategies within the context of formal instruction. In the case of self-instruction in the absence of an interlocutor, these strategies become even more pertinent. Such "less authentic" uses of PL need not (and indeed cannot) substitute listening practice of a communicative nature. However, they are readily available to many learners for whom they may well appear to be the most natural option. It would seem a shame to dismiss such strategies before we even know their potential effects on learning.

The discussion of PL has also brought to attention two other matters. Firstly, that the degree of orality in a recording (e.g. scripted/semi-scripted/unscripted) is an essential feature of listening input, all too often overlooked in listening research. The other feature that the discussion of PL has made particularly apparent is the temporal dimension of listening. A few time-related variables have already received considerable attention from researchers (e.g speech rate, effect of pauses etc.). Other temporal aspects, such as the sequence of information presented in PL (e.g. effects of repetition on "non-redundant" ideas, location of a segment at the start or the end of a recording) and its effects on listener behaviour deserve greater attention.

Sequence-related aspects of processing and strategy use in PL are also worth investigating, as seen from the results of experiments 2 and 3.

10.2 Methodological advances from PL research

The second question formulated at the beginning of this study was:

- *How can PL be used for research on listening processes and strategies?*

A number of research instruments were specifically developed for the purpose of examining PL related data. They cover two different areas: instruments for the analysis of listening input used in PL tasks (discussed in section 6.1: "Operationalising meaning-related variables in ARI"); and instruments for the analysis of listeners' interaction with the listening input (discussed in section 6.2: "Operationalising strategy-related variables in PL").

10.2.1 Idea-maps

Some of the tools for analysing the input, such as the identification of idea-units (Chafe 1985; Crookes 1990; Kroll 1977; Meyer 1975) and the division of recordings into segments (Blau 1990; Merlet and Gaonac'h 1995), had already been widely used in listening research. However Experiment 1 needed to establish the precise segment(s) of the recording from which specific idea-units could be either decoded or inferred. Idea-maps were developed for this purpose (see p.121).

Idea-maps have several advantages as research tools:

- they locate semantic information within the temporal sequence of speech
- they make it possible to quantify the level of reduncancy of a given idea-unit according to the particular playback sequence followed by whoever operates the playback equipment

One important limitation of idea-maps was also identified in Experiment 1: it is not always possible to pinpoint an idea-unit to one (or several) specific segment(s) of the input. Recordings in which a large amount of inferencing is required from the listener may not lend themselves to being represented in the form of idea-maps.

The method is best used for the study of recordings requiring types of processing to which Brown refers as "identifying information" and "procedural understanding", i.e.

"which may only require a minimum of linguistic interpretation" (Brown 1994)[38]. In many instructional settings, listening tasks tend to follow a traditional routine in which listeners are prompted to retrieve the propositional meaning of specific segments through a series of sequentially organised "comprehension" questions (Brown and Yule 1983b:56). The target idea-units in such cases should be relatively easy to map. However, idea-mapping would become problematic if semantically complex recordings or open-ended tasks such as free recall were used instead. Given that the kind of listening task described above is still in use in many teaching environments, idea-maps should have a place in L2 listening research.

10.2.2 Move logs

The second type of instrument developed in this study was designed to assist in the analysis of listeners' interaction with the PL input. The fundamental tool for this purpose is the move log (see p.125). It enables researchers to produce a graphic representation of the sequence of playback operations performed by the person controlling the equipment (in this study, listeners themselves). The vertical axis of a move log represents the chronological sequence of *segments* in the recording, whereas the horizontal axis shows the successive *moves* made by a particular listener. Unlike idea-maps, move logs can be easily applied to any type of recorded input.

Only three weaknesses were identified when assessing move logs as research instruments:

(1) The input needs to be arbitrarily segmented by the researcher in one way or another. Although the method used in this study produced fairly short segments, it can still be argued that there is a loss of information when a subject pauses the tape half-way through a segment.

(2) Listeners are not necessarily accurate in their use of playback controls (e.g. rewinding further than intended). Indeed, a few moves were identified as possible errors in the three experiments. Audio cassettes are problematic in this respect. The use of a medium allowing pre-defined segments (such as the tracks on a CD) would alleviate this particular problem, but would aggravate problem (1).

(3) Most importantly, move logs tell us what listeners do with the input, but they do not tell us why they do it (cf. Experiment 2). Therefore, triangulation with a verbal

[38] See p.51.

method of some kind (cf. Experiment 3) is indispensable for a full understanding of the processes and strategies involved in PL.

Using the set of 32 move logs that was generated by 16 listeners while performing the two listening tasks presented in Experiment 2, a series of general playback patterns were identified and named (see section 8.3.4). Each pattern was hypothetically related to certain PL processes and strategies, and these hypothetical relations were subsequently tested against move logs and verbal information from the listeners in Experiment 3.

10.2.3 Paired listening protocols

To address the difficulty of interpreting playback behaviour on the basis of move logs alone, Experiment 3 involved a combination of move log data and data from verbal protocols. In order to prevent the shortcomings of introspective reports (Matsumoto 1993; 1994; Seliger 1983), verbal protocols were not based on think-aloud techniques, but rather on the verbal interaction between peer listeners during paired listening tasks. The original reservations about the use of introspective methods were confirmed in Experiment 1, where a few subjects supplied misleading reports in response to the question: "Did you pause the tape very often?" (see Table 7-2, chapter 7).

Paired listening protocols offer the following advantages:

(1) The purpose of talking is to complete the task, not to pass on information for research purposes. Strategy information is analysed directly by the researcher and subjects are not required to mediate through introspection.

(2) Focus can be centered on overt behaviours and observational data: encoding conventions of specific behaviour are explicitly formulated (e.g. "L1 word", "L2 chunk", "global inference", see section 6.2.2.2).

(3) There is no time lapse between performance and verbalisation: data collection is concomittant to the task, not retrospective. This prevents loss of information once it is no longer held in STM.

(4) Students work in a familiar language laboratory situation that could occur in one of their language lessons. This is the environment in which foreign language learners are most likely to encounter PL in real life. This implies that the findings are more likely to reflect naturally occurring phenomena than findings obtained in other experimental settings.

(5) Without the physical presence of a researcher, learners soon forget that they are being recorded, and their social-affective strategies are more evident than they would be with conventional introspective methods. The task partner also acts as a powerful stimulus for this type of behaviour.

(6) Talking aloud by yourself while performing a task in front of a stranger is an artificial form of communication; talking to a task partner makes sense because communication is a necessity.

There are, however a few limitations as well:

(1) Individual variables such as language proficiency are difficult to control when the task is performed in pairs.

(2) The strategies used in paired listening are not necessarily identical to those used by listeners when they are working by themselves on recorded input. At this early stage, so little is known about specific PL strategies that such differences are probably of little relevance in comparison to what can be learned from any existing similarities. At some point however, it will be necessary to examine the two modes of PL as different categories.

(3) The main purpose of talking is to complete the task in hand. This is an advantage insofar as it does not require the subjects to interpret their own behaviour (Seliger 1983), but it also leaves the researcher with many unanswered questions. Some playback behaviours can be explained through an analysis of verbal information from the partners' interaction, but a large proportion still remains unaccounted for. Once the full potential of observational methods has been exhausted, introspection could still be the only means of obtaining certain types of specific information related to PL. For instance, one could follow up paired listening by playing back the conversation separately to each of the the participants, prompting introspective commentaries as appropriate.

(4) Considerable effort was made to design a set of coding conventions based as tightly as possible on specific observable behaviours. Unfortunately, some of the categories proposed (e.g. inferencing) can be difficult to classify using strictly objective criteria. Here again, the criteria used in this study could be improved (for instance through the systematic introduction of inter-rater reliability tests). Given the inevitable time constraints, greater critical depth in this area would have affected the breadth of the sample and task types that were used in the present study. The diversity of PL strategies found, and their dependence on input and

task variables confirm that such a wide coverage was essential at this initial stage. However the need to eliminate any equivocal criteria from the analysis instrument must be addressed eventually. This is particularly necessary if paired listening protocols are to be used for investigating more specific aspects of PL.

In spite of the limitations discussed above, the combined use of idea-maps, move logs and paired listening protocols to examine PL strategies and processes opens up a much needed observational dimension in listening strategy research. The next section evaluates the preliminary findings obtained in this study and outlines a number of promising directions for future research.

10.3 L2 listening processes and strategies in PL

The third question formulated at the beginning of this study was:

- *What does initial PL research tell us about L2 learners' listening processes and strategies?*

On the basis of the three experiments conducted, a number of relations between playback strategies and listening processes/strategies can be established. These will now be examined in turn. The discussion below will show that, in many cases, rigorous one-to-one relations between playback strategies and listening processes/strategies would be inappropriate. The reported behaviours were found to involve a variety of inter-related factors: often a similar playback pattern occurring in different move logs was found to be related to a number of distinct processes. The factors determining the different relations found are thought to derive from specific combinations of input, task, and individual processing ability. Consequently, certain playback strategies will be mentioned under more than one heading.

10.3.1 Pace of progression and cognitive processing

The amount of input that subjects were able to tackle by means of automatic processing was found to play a key role in determining the pace of progression through a recording. A slow pace (few segments played at a time) indicates that controlled processing is taking place, whereas a fast pace (many segments in one go) means either that processing is largely automatic, or on the contrary that no processing at all is taking place.

One of the characteristic features of controlled processing is that it takes significantly longer than automatic processing (Schneider and Shiffrin 1977; Shiffrin and

Schneider 1977) and only "fluent" subjects who can process automatically are able to perform several processes at a time (Just and Carpenter 1992; Rabinowitz and Chi 1987; Schmidt 1992). Therefore, as Merlet and Gaonac'h (1995) hypothesised, controlled processing of listening input proceeds in small chunks, often going back to the same segments again, so that the input can be processed one little step further. This, however, is only worth doing when a minimum amount of processing is possible in the first place. Below that minimum, subjects just give up and let the recording play on. This is consistent with Blau's findings on the effects of pauses on comprehension (Blau 1990), where pauses were found to be most useful when difficulty was neither too high or too low for the listeners. The results of Experiment 1 in this study also showed that fewer pauses were made by subjects in the easiest recording. Most pauses were made in the other two (more demanding) recordings, with the exception of one very difficult passage, in which pauses were again very rare. Retrospective comments made by subjects in Experiment 1 confirmed this relationship between the number of pauses made and input difficulty.

A potentially interesting project for further investigation would be to establish the parameters that define the threshold and ceiling levels for a slow progression to occur. Presumably, a recording in which L2 learners use controlled processing should be at an ideal level for learning to take place. Once processing becomes largely automatic, listeners are probably operating below their zone of proximal development (ZPD) and are no longer acquiring new processing skills (though it could be argued that the existing connections continue being reinforced through automatised practice). If, on the other hand, the progression is fast because listeners are unable to process anything, the input is clearly beyond their ZPD and no learning is possible at all. Since the ideal level of difficulty lies within the ZPD (de Guerrero and Villamil 2000; Hossein and Cumming 2000; Vygotsky 1978), the pace of progression could be used as an indicator of the relative difficulty of a recording and its suitability for a particular subject at a given stage of L2 development.

10.3.2 Patterns of variation in playback frequency

A tendency to play the beginning of a recording more times than the end was found in most of the recordings used. The only case where this tendency was not observed was the easiest of the recordings, whereas the difference in frequency was most evident for difficult recordings. This phenomenon could either take the form of a clearly marked difference between the first and second halves of a recording (a

playback pattern that was termed "top-heavy"), or that of a more even, gradual decrease in playback frequency throughout the entire recording (termed "gradual decrease").

The findings suggest that the two patterns might be related to different processes: distinct top-heavy patterns appear to occur when subjects experience difficulty in finding a valid schema on which to base a global meaning hypothesis for the recording. When the information available is not sufficient to activate this type of initial schema, subjects concentrate their search for further clues on the opening segments, at the expense of later sections of the recording.

In cases where playback frequency decreases more gradually, subjects appear to have reached some kind of global schema already, however basic or inaccurate. In such cases, the gradual decrease in playback frequency is roughly proportional to the overall difficulty of the recording. The more difficult a recording is, the steeper the decline in playback frequency towards the end.

One hypothetical explanation for the most distinct top-heavy patterns could be that subjects assume that key information to the global schema of a recording lies in its opening segments (a common assumption that proved wrong for the example used in this study: recording D). Further studies eliciting more explicit responses from the subjects would be required to confirm or disprove this hypothesis. Provisionally, it is suggested that the learners' assumption probably operates in combination with other phenomena, some of which might be the same as those accounting for more gradual decreases in playback frequency.

Gradual decreases can be interpreted in terms of the combined effects of primacy and fatigue: subjects concentrate their efforts on the first segments presented to them. By the time they reach the end of the recording, fatigue has gradually set in, causing the amount of controlled processing undertaken to decrease accordingly. The greater the effort required at the beginning, the more fatigue subjects will experience, and the more likely they will be to give up playing segments repeatedly later on.

Just and Carpenter (Just and Carpenter 1992) argue that the need to maintain several hypotheses active simultaneously causes strain on STM capacity. The authors even quote evidence from observed physiological effects (such as glucose metabolism) related to the effort made by less skilled subjects in problem solving tasks (Haier et al. 1988, quoted in Just and Carpenter 1992:143). When input is easy,

one would consequently assume that subjects are more able to maintain an even level of activation throughout the recording, thus engaging in comparable levels of processing activity from beginning to end. In contrast, difficult input requires a higher proportion of controlled processing (reflected in more repetition of the same segments), heavier STM load, and less ability to maintain all the required connections active from beginning to end. As most subjects proceed sequentially through this kind of task (see below), they will eventually lose track of some connections and experience a gradual decay in their ability to keep up controlled processing. Consequently the pace will gradually become faster (using longer moves) as a result of growing lack of processing as they get closer to the end.

An even simpler phenomenon could explain the correlation between decrease in playback frequency and task difficulty: when processing is delayed or interrupted by a difficult passage, subjects prefer to rewind back to the beginning and "try again", possibly in an attempt to find a connection between the problematic section and any information that they have already got. This results in the beginning being played more often than the end. The effect is also likely to be more pronounced when more problems are encountered along the way, hence the correlation with task difficulty.

The first aspect that needs further investigation is whether "top-heavy" and "gradual decrease" patterns are just slightly different instances of the same phenomenon, or whether they reflect two clearly distinct processes. The hypothesis whereby top-heavy patterns indicate the lack of a valid initial schema could be tested by exposing different groups of subjects to identical input and controlling the levels of schema activation by means of different advance organisers (supplying a title for the recording / preview comprehension questions / no clues given at all). The "primacy + fatigue" hypothesis could be tested by varying the duration of otherwise identical input. For instance, the first group would do a free playback listening task using recording A (say 4 minutes long), the second group would work on recording B (= the first 2 minutes of recording A), and the third group would work on recording C (= the last 2 minutes of recording A). The study of move logs (and possibly paired listening protocols) in both experiments ought to help us understand what exactly is causing such differences in playback frequency.

10.3.3 Sequential approaches to playback strategy

In most cases, subjects were found to follow a sequential approach to the tasks set. This was particularly evident in Experiment 2: out of the 16 subjects who were

presented with two listening tasks, only one took the trouble to listen to both recordings before starting work on the first one. All the others spent several moves on the first recording before even listening to the second recording once.

In both Experiment 2 and Experiment 3, numerous instances of progressions were found in which the clusters occurred in sequential progression (indicated on the move logs by clusters arranged along a diagonal line descending from left to right). In other words, they focused on successive groups of segments, following the sequence in which information was presented in the recording. The strategy of proceeding in strict sequence through the recordings might be driven by the underlying assumption that, if you make sense of the beginning, then the next bit will become clear, and then the bit after that, and so on. On the other hand, subjects might be operating merely by intertia, using the same type of sequential approach which led 15 out of 16 subjects to begin work on the first recording, before even trying to find out what the second recording was like.

An interesting case was recording F, in which comprehension questions followed a slightly different order ("question sequence") from the sequence of information in the recording ("text sequence"). Protocol analysis results show that playback strategies were normally determined by a combination of text sequence, question sequence, and incidental recognition of salient items through automatic processing. The resulting patterns found in move logs were hardly neat reflections of one sequence or the other. However, protocol data did show just how much relevance such sequences (and any existing discrepancies between the two) could have to the metacognitive strategies that subjects adopted through the task.

Since the first two experiments in this study were exploratory and hypothesis-generating, the order of questions was not originally controlled. The initial concern in Experiment 2 was that students were to operate in as natural conditions as possible, hence the inclusion of a familiar test format using comprehension questions. Only a closer study of the resulting move logs revealed the potential relevance of cluster location to strategies related to text/question sequence. Given the complex interactions later found in protocol analysis, a further experiment would now be required, using a task in which text and question sequences should be unequivocally different from each other (this was not the case for recording F). A second group of subjects would listen to the same recording, this time with the questions presented in the same sequence as the text, while the control group would perform a recall task without the aid of questions. Further experimental designs to compare sequence-

related strategies in listening tasks using "static", "dynamic" and "abstract" types of input (Brown and Yule 1983b) could also be considered.

10.3.4 Cognitive processing in the most repeated segments

A potentially important finding of Experiment 2 was that the most frequently played segments were also the best recalled. Further examination of protocol data in Experiment 3 showed that in every single instance of high repeat-rate (i.e. when RR was at least 1 standard deviation above average), some amount of information related to the repeated segments had been decoded in subsequent hearings. This confirms the hypothesis whereby repetition does increase the occurrence of bottom-up decoding processes (Merlet and Gaonac'h 1995). Inferencing strategies were also found to occur as a result of repetition, but the effect was not universal and appeared to depend on input, task and individual variables.

The hypothesis whereby the most repeated segments are also those in which subjects recognised items in the first place was only confirmed for the task in which subjects were listening in order to answer a list of comprehension questions (recording F).

This implies that repetition strategies are universally beneficial in terms of decoding strategies (albeit in varying degrees), but they may also be related to the strategies of inferencing and listening for specific information if certain individual, input-related and task-related conditions are met.

10.3.5 Strategic stages of playback procedure

The original prediction was of an "a-b-a" playback pattern in three stages (fast-slow-fast). Three hypothetical strategies were associated with the initial fast progression: getting acquainted with the speaker's accent and delivery, locating relevant information and activating relevant schemata. The second, slow progression stage was associated with intensive bottom-up processing of the input. The final stage predicted consisted of a fast progression, associated with hypothesis-checking strategies.

Only two clear occurrences of this playback pattern were actually found in the move logs (one subject in Experiment 2 and one pair in Experiment 3). Protocol analysis based on the single case found in Experiment 3 showed that the pair who adopted

this playback pattern was also the only one who successfully understood the main topic of their (very difficult) recording. It was therefore concluded that although the a-b-a strategy is far from common, it would appear to be highly effective. The pair introduced a few minor changes in the strategies hypothetically associated with the different stages, but these are believed to be adaptations to the specific constraints of the task in hand.

A closer study of the strategies observed in the first and last runs across the whole sample in Experiment 3 confirmed the partial validity of some of the relations predicted in the a-b-a hypothesis:

All three strategies associated with stage 1 of the a-b-a strategy (the first, "fast" run) were actually found to occur in several instances across the sample, whether they proceeded in fast or slow progression. Certain conditions seem to determine which particular combination of such strategies is used in each case: reactions to the input are more frequent when the input presents some difficulty (no such reactions were found in the easiest recording); overt schema activation is not used if a schema is readily available (either from questions preview or from immediate automatic decoding of key information); locating information is only possible if a schema is already in place to indicate which information slots are to be expected, and provided that the relevant information has not been automatically decoded already. An additional strategy often found at the end of the first run was evaluating the difficulty of the input and/or one's own performance. This strategy had not been predicted in the original hypothesis.

As for the final run in a task, numerous instances of hypothesis-checking strategies were found, although once again, they occurred in both fast and slow progressions. In cases where large amounts of controlled processing had been required, a complete recap of the final meaning hypothesis also took place after the last run.

In conclusion, the emphasis that was originally placed on the pace of progression in opening and closing runs proved to be irrelevant. Nevertheless, the predicted relations between the initial and final runs of a PL task and their associated listening strategies were largely accurate, with only a few variations determined by input and task-specific variables.

Further studies would need to determine more precisely in what specific conditions each possible combination of the strategies above is most likely to be used. For this

purpose, a fairly comprehensive battery of paired listening experiments would be required. These would involve a range of recordings carefully analysed in terms of input difficulty, and presented with different levels of schema activation (e.g. preview title, preview questions, no clues, etc.).

The effectiveness of the a-b-a strategy is also worth testing: if its benefits were confirmed, it could be introduced as part of standard strategy training in L2 listening instruction. Given the rare occurrences of a-b-a playback patterns in "free playback" experimental settings, it would be necessary to induce the strategy either by instructing subjects to use it, or by returning control of the playback procedure to the researcher. Thus, the sample available would be much larger and could also include a wide proficiency range across the subjects. In the researcher-controlled setting, a control group could be exposed to an identical number of repetitions as the "a-b-a" group, but with the recording presented exclusively in straight runs instead.

10.3.6 Social/affective strategies in playback listening

The most significant finding in relation to social/affective strategies was the sheer amount of information elicited by the paired playback listening method used in Experiment 3. Much of the data obtained (laughter, hedging strategies, overt sighing, deliberately far-fetched hypotheses, etc.) was not systematically analysed due to its lack of direct relevance to playback strategies.

Three types of relationship between task partners were identified: balanced cooperative, unbalanced cooperative, and uncooperative. In terms of specific relations to playback listening strategies, balanced cooperative partnerships tended to result in slow progressions, whereas uncooperative partnerships resulted in fast progressions.

De Guerrero and Villamil (2000) argue that, when two task partners are working cooperatively, they tune their strategies in order to operate within the ZPD of the partner with lowest proficiency. In section 10.3.1 above, it was suggested that the typical playback strategy used when operating at the ZPD would be a slow pace of progression, given that a large amount of controlled processing was taking place. This playback pattern was observed in all cases of balanced cooperative partnerships found in Experiment 3. However, it was not the case in the single occurrence of clearly unbalanced cooperation found in the sample, where the pace was relatively fast. A possible reason for this could be that the most proficient subject was speeding up the pace by providing ready-decoded segments to his partner. Subjects who are

better able to process the input automatically than their partners may spare them the need for repeated exposure, thus causing the pace of progression to be faster.

In such instances of one-sided recognition, another cooperative strategy was often observed, in which more proficient subjects would supply the ready-decoded segment to their partners, and then locate and play back the relevant section of the recording, to give their less proficient partners a chance of backtracking the process themselves. This strategy, which is a clear example of scaffolding, confirms the claim that cooperative task partners tune their strategies to operate within the lowest common ZPD. It should also be noted that this type of cooperative strategy has the effect of slowing down the pace of progression.

The case of non-cooperative partnerships deserves particular attention. Two cases were found in the sample. They both produced fast progressions, and protocol data showed unilateral decisions to terminate the task prematurely in both cases. The reasons however appeared to be slightly different for the two pairs. In the first case, the non-cooperative member of the dyad displayed a clearly hostile attitude to the task. His mention of previous negative listening experiences suggests that he hastily "decided" that the input was too difficult to process and therefore not worth the effort (one type of fast progression occurs when no processing is possible). In the second case, the subject's attitude to the task was positive, yet he seemed in a hurry to see the job completed. As a result, the very first, hasty hypothesis was adopted without ever being checked. This caused the pair to settle for a highly inaccurate global interpretation with virtually no inclusion of specific information. Here again, lack of processing resulted in fast progression.

The tendency of low proficiency listeners to settle for a premature, unchallenged hypothesis could be explained in terms of the inability to maintain parallel hypotheses due to limited STM capacity. In a study mentioned earlier, Just and Carpenter (1992) argue that maintaining simultaneous hypotheses causes excessive strain on STM capacity when proficiency is low. The combined effects of greater effort and lower levels of activation cause the less obvious hypotheses to decay in favour of the ones requiring least effort. High proficiency students on the other hand are able to maintain several hypotheses active in parallel until one of them is unequivocally confirmed. The processes examined in the Just and Carpenter study were based on largely automatised, unconscious behaviours (e.g. gaze durations measured in milliseconds). When the processes involved are of a controlled nature (such as in the decision to rewind and listen to a passage again), the effects are likely to be even more evident,

since controlled processing requires attentional effort. It is therefore not surprising that low proficiency subjects give up and settle for unsatisfactory interpretations more readily than their proficient counterparts.

Refusal or reluctance to process input is also believed to be a common problem among insecure L2 learners, yet little research has been carried out on this type of avoidance strategy. Seliger (1977) observed the amount of verbal interaction in which learners engaged over two hours of language class, as well as the proportion of verbal interaction that was actually initiated by learners themselves. He identified two types of language learners: high input generators (HIG) and low input generators (LIG). Seliger proposed that, by initiating interaction, HIGs cause responses that result in greater exposure to L2 input, whereas LIGs adopt a merely reactive behaviour in which they interact only when prompted by the teacher. These two behaviours could be compared respectively to those observed in cooperative pairs (who exposed themselves to large amounts of input by making a large number of moves) and non-cooperative pairs (who simply "reacted" to the task brief by playing the tape only a few times, thus depriving themselves of the opportunity to process the input as much as time would have allowed). Given that high levels of repetition have been found to be consistently related to subsequent decoding processes, it could be inferred that LIGs are losing out by not playing the recordings as many times as they could.

Finally, a number of specific strategies were found to occur at particular stages of the listening task:

Whenever overt affective reactions to the input were found (normally in the form of negative reactions when input was perceived as "difficult"), these tended to occur immediately after (or even during) the first run. In other words, before explicit cognitive processing began. In contrast, evaluations of the subject's own performance did not occur during the intervening moves of a run, but at the very end of the last move made in the run. Presumably, while moves are being made as part of controlled or automatic processing, subjects' attention is entirely focused on cognitive processing and cannot be spared for metacognitive/affective strategies such as self-evaluation. Only once the end of the recording has been reached can they spare attentional resources to perform an explicit evaluation. This may occur either once the task has been fully completed, or at an intermediate stage, when a run has just finished and a new one is about to begin.

One important conclusion of this discussion is that more research is needed in the area of "negative" affective strategies. When strategy research mentions social/affective strategies, it generally refers to facilitating strategies such as self-encouragement. Yet as this study shows, a substantial proportion of affective strategies is of a reactive or adverse nature. This is not to say that all negative statements should necessarily carry adverse effects: a few such behaviours could just be face-saving strategies intended to facilitate the rapport with an equally insecure task partner, or healthy means of releasing tension after the stress of on-line processing. In any case, a better knowledge of the nature, timing and function of negative strategies would significantly contribute to our understanding of the real function of affective variables in L2 listening strategies.

10.4 Recommendations for the teaching and testing of listening skills

In the light of the previous discussion, a number of issues are of relevance to the final question that this study aimed to answer:

- *How can PL be used in order to develop listening skills in L2 ?*

The results obtained in the present study suggest a number of recommendations for course designers, language teachers, test designers and language advisors. They should be read within the perspective of Rees-Miller's (1993) cautious advice against over-prescriptive approaches to learner training (see p.104).

10.4.1 Pace of progression

The pace of progression can be used as an indicator of input difficulty. Ideally, teachers and course designers should aim for recordings that produce a slow pace of progression, because this means that learners are operating within their ZDP and learning potential is at its maximum level. This approach can be used for materials selection purposes, and also as a learning exercise at the "controlled practice" stage (Littlewood 1981).

Fast progression also has its place in a listening syllabus, so long as it is related to a high proportion of automatic processing. Such an approach would be appropriate for extensive listening activities, by reinforcing existing connections through practice. In the long term, this would improve performance in more demanding tasks, because

stronger connections mean greater processing capacity (Just and Carpenter 1992). Automatic processing is also a sign of success and should have beneficial confidence-building effects that could reduce listening anxiety overall (Vogely 1998). A fast progression could also be appropriate when the listening task is one in a series of tasks within an integrative learning activity covering a variety of skills. Fast progressions elicit modes of processing that are closer to those typically required in real-life situations. However, if a fast progression is due to lack of processing, either the input or the task variables need to be changed so that at least some amount of processing becomes possible. Failing this, the listening task that generated such a fast progression would be of no benefit to the learner.

10.4.2 Playback frequency and global playback patterns

Most repeated segments are always associated with higher recall rates and further decoding processes. Therefore learners should be encouraged to focus on all sections of the input, and not only the beginning.

If found to be effective in further playback strategy studies, the a-b-a strategy could be taught explicitly as part of the listening curriculum. Its introduction could be particularly beneficial to counteract the overuse of top-heavy patterns, which distract learners from finding the clues that may be available in later sections of the recordings.

Low input generators should be encouraged to become high input generators: fast progressions due to a "no point" approach can be prevented by avoiding exposure to very difficult input in which no processing at all is possible for the learners. When recordings are difficult, a wide range of scaffolding strategies can be used (advance organisers, glossary, gapped summary, etc) to support low-proficiency listeners and make some amount of processing possible. Observational indicators such as task completion time and pace of progression can be monitored in order to adjust the amount of assistance required in each instance.

10.4.3 Strategic sequence in free playback listening tasks

The widespread tendency among the subjects to proceed in linear sequence through listening tasks needs to be taken into account when designing listening tests. Most test designers already tend to put the easiest tasks first, grading the tasks by increasing difficulty (e.g. GCSE, see Oxford Cambridge and RSA Examinations (OCR) 2000c). This intuitive approach seems to be justified. By increasing the

processing load gradually as the task progresses, subjects are likely to use more automatic processing at the beginning, saving up their energy for the more difficult tasks that require a greater proportion of controlled processing. This is particularly important if test takers are given control of the tape, since a difficult first recording could elicit slow progressions that would use up too much time before subjects even attempt the remaining tasks.

In listening tests that allow free playback, only a minute proportion of the sample (1 out of 16) was found to use the strategy of listening to all the recordings given in a test before tackling the first one. By restricting their attention to the first recording in the sequence, subjects are denying themselves the opportunity to plan their playback strategy according to task and input demands. Specific training to prevent this could be useful. From a test designer's point of view, a free-playback listening test ought to include directions on the amount of time recommended to be spent on each task, or even direct advice on specific playback strategies (e.g. "Listen to *all* the recordings once before you answer section A"; "Make sure you spend *no more than 20 minutes* on each recording").

Another important issue in test design is the sequence in which comprehension questions are presented. A sequence that does not conflict with the sequence of information in the recording would be easier to follow. On the other hand, this approach has been criticised by some authors (Brown and Yule 1983b; Schrafnagl and Cameron 1988), on the grounds that it does not elicit the kind of listening strategies that we use in real life. If the preferred question format involves scanning for specific information in a sequence that is either not consistent with the text or with the cues given (e.g. questions, diagrams, cloze summaries, etc.), the strategies required in order to obtain the necessary information will be more complex. Test designers and teachers need to be aware of this added complexity and be prepared to provide adequate training beforehand.

10.4.4 Paired listening with free playback

Paired listening was used here as a data collection method on free playback strategies, but the levels of processing and interaction found in the protocols indicate that it could also be a very powerful teaching method. The only precaution to be taken would be to prevent the occurrence of non-cooperative partnerships. A number of measures could be taken to ensure a cooperative attitude:

One should make sure that the task is not so difficult that it promotes a "no point" approach to processing: whenever the input presents significant difficulty to the target learners, task requirements ought to be adjusted accordingly so that adequate support features are included (glossaries, advance organisers, etc.). As seen in the paired listening protocols of Experiment 3, one frustrated student can soon make the other give up too.

When pairing up students, one should aim to match students of similar proficiency together. If some differences in proficiency are inevitable, unbalanced matches in which the lowest proficiency student is the most assertive in the pair ought to be avoided.

Cases of early task completion should be monitored in order to assess the causes for fast progression. If excessive or insufficient difficulty was found to be the cause, the task could be adjusted accordingly (for instance, by giving students additional questions that made it either easier or more challenging); if the problem was more related to pair dynamics, different matches could be arranged to achieve a more productive balance.

The maximum benefits of paired listening are believed to be gained when the combination of input, task, and individual variables causes learners to proceed in reasonably slow progressions, with a fair amount of controlled processing, ideally resulting in final hypotheses that have been appropriately checked. Brown and Yule's concept of "reasonable interpretation" (Brown and Yule 1983b) as a measure of success is particularly relevant here.

Section 10.4.1 above also mentions cases in which fast progressions would be appropriate, and the use of paired listening for such activities could be considered as well. However the data available in the present study cannot be used to assess the potential benefits of paired listening in tasks that elicit fast progressions.

10.4.5 "Adverse" affective reactions

More needs to be understood before a valid set of informed recommendations can be made in this area. It could well be that voicing negative reactions to the input is actually a necessary step in playback listening. One aspect that may be relevant to PL procedure is the spontaneous timing of such negative reactions. The findings of Experiment 3 seem to suggest that cognitive processing does not normally begin until the initial affective reactions have been expressed. Teachers should be aware of this

and give the learners space to react affectively before being asked to process the information. It might even be appropriate to prompt reactions to the input in a controlled manner at the beginning of a task (e.g. asking "How did that sound then?" after the first run). The teacher would then be able to assess the learners' state of mind and if necessary, prevent anxiety or frustration from setting in. A simple acknowledgement of the initial impact of fast speech rate could be all it takes to alleviate the stress in a majority of cases.

Even when the on-line processing constraints are reduced through the use of repeated playback, the object of listening remains intangible and elusive because of its very nature: sound is time-bound and invisible (this could be yet another argument in support of video recordings). Anxiety is always a potentially important factor for L2 listeners. Yet there is surprisingly little research in this area, in comparison to other areas of language anxiety such as speaking anxiety (Vogely 1998). Here again, the potential benefits of paired listening could be significant: sharing responsiblity for the successful completion of the task could ease the pressure on single individuals. Self- and peer-encouragement have often been mentioned as effective strategies, but perhaps other affective strategies of a less "positive" nature also have a role to play. Gaining a peer's approval on the perceived "unreasonable" demands of a given task is a way of legitimising one's own limitations. Provided that such shared negative perception of the input is not allowed to dominate the task, it might actually help the learners relax and accept the levels of uncertainty that any L2 learner has to learn to live with (Brown and Yule 1983b; Rubin 1975). The teacher's role is to be vigilant, so that truly negative self-perceptions are not allowed to develop. Overall, the aim should be to promote a gradual improvement of the learners' perception of their own performance through regular success on reasonably demanding tasks. Learners themselves have voiced the request to be given opportunities to experience "small successes" (Vogely 1998). Hence the importance of careful grading and monitoring.

10.5 Summary of conclusions

The present study represents an initial investigation of the notion of playback control as a relevant feature in a listening task. The overt nature of playback behaviour was exploited in order to propose a series of research instruments -such as move logs- based on the notion of PL. Verbal data from paired listening protocols were also used to assist the interpretation of playback behaviour. The results of this process show

that the methods proposed can indeed contribute to the advancement of our knowledge about listening processes and strategies, especially those related to listening tasks involving recorded material. The methodological limitations of this approach were also discussed, but the nature and quantity of information that they are hoped to supply was considered more significant than any possible drawbacks at this stage.

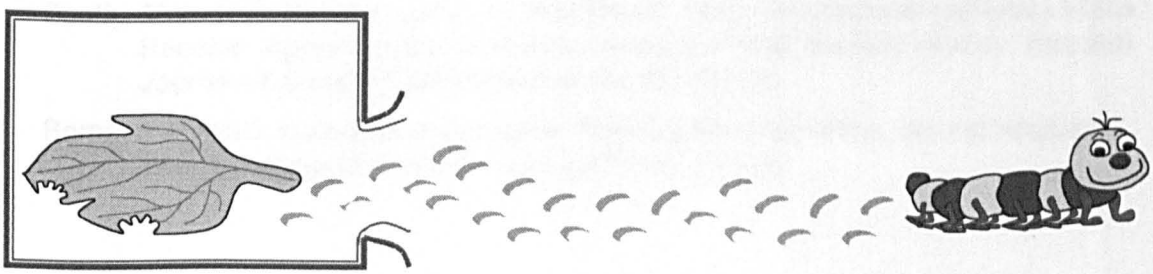
In terms of results, this study established a number of specific relations between playback strategies and listening processes/strategies. Firstly, the pace of progression in a PL task indicates the type of processing that is taking place: a low progression indicates controlled processing and (in paired listening tasks) a balanced cooperative partnership, whereas a fast progression indicates that processing is either automatic (with 'easy' input) or inactive (with 'difficult' input), and is normally related to non-cooperative or unbalanced partnerships in paired listening tasks. Secondly, subjects appear to have a tendency to play the beginning of the recordings more than the end. The decrease in playback frequency is generally gradual, and most evident when the input is demanding on the subject's ability. If the listener struggles to find an initial schema on which to base subsequent meaning hypotheses, the initial segments are played again and again, producing distinctive top-heavy playback patterns. In general, the most repeated segments are also the best recalled, and intensive repetition promotes bottom-up decoding. Finally, it was found that subjects typically approach PL tasks in a sequential manner and that the sequence of questions in a test interacts with the sequence of information in the text and determines the playback strategy used.

For the language teacher, the main implication of this study is that PL is a valid type of learning activity. The skills and strategies involved in PL are not (and need not be) the same as those used in real life situations outside the instructional environment. However, such PL skills and strategies are pertinent in the context of formal instruction and self-study situations. Until more is known about PL, this mode of listening should not be dismissed on the grounds that it does not develop 'authentic' skills. The preliminary results obtained in this study suggest that a few recommendations can already be made at this stage.

Analysis of playback strategies can help teachers diagnose the relative difficulty of a recording for a particular group of learners and the kind of processing in which listeners engage. Recordings generating slow progressions can be used for improving controlled processing (especially bottom-up decoding), whereas those

generating fast progressions can be used for reinforcing existing skills, confidence-building and integrative listening tasks that approach those encountered by listeners in real life. Paired listening is also recommended as a potentially useful learning activity. Finally, a number of affective responses related to PL were observed, particularly in the form of initial negative reactions to the input (especially to speech rate) before cognitive processing begins in a listening task. Until more is known about the function of such initial reactions, teachers are advised to allow these to occur, as they might serve the purpose of releasing tension before commencing a cognitively demanding task .

Much remains to be learned from the study of PL, but the results outlined above suggest that it is a promising area for future research. It is also hoped that strategy research will benefit from adding the observational approach suggested in this thesis to the variety of introspective methods that are currently in use.



11. References

Allan, M. 1985. *Teaching English with Video, Longman Handbooks for Language Teachers,*. London: Longman.

Altmann, G.T.M., A. Garnham, and Y. Dennis. 1992. Avoiding the garden path. Eye movements in context. *Journal of Memory and Language* 31:685-712.

Anderson, A., and T. Lynch. 1988. *Listening.* Oxford: Oxford University Press.

Anderson, J.R. 1976. *Language, Memory and Thought.* Hillsdale, N.J.: Erlbaum.

Anderson, J.R. 1982. Acquisition of Cognitive Skill. *Psychological Review* 89:369-406.

Anderson, J.R. 1985. *Cognitive Psychology and its Implications.* New York: Freeman.

Anderson, J.R. 1989. Practice, working memory, and the ACT* theory of skill acquisition: A comment on Carlson, Sullivan and Schneider. *Journal of Experimental Psychology: Learning, Memory and Cognition* 15:527-30.

Anderson, J.R. 2000. *Cognitive Psychology and its Implications.* 5th ed. New York: Worth Publishers.

Anderson, R.C., and P.D. Pearson. 1988. A schema-theoretic view of basic processes in reading comprehension. In *Interactive Approaches to Second Language Reading*, edited by P. L. Carrell, J. Devine and D. E. Eskey. Cambridge: Cambridge University Press.

Atkinson, R.C., and R.M. Shiffrin. 1968. Human memory: A proposed system and its control processes. In *The Psychology of Learning and Motivation*, edited by K. W. Spence and J. T. Spence. Orlando, FL: Academic Press.

Bacon. 1992a. Authentic listening in Spanish: how learners adjust their strategies to the difficulty of the input. *Hispania* 75 (1):398-412.

Bacon, S.M. 1992b. Phases of listening to authentic input in Spanish. A descriptive study. *Foreign Language Annals* 25 (4):317-34.

Bacon, S.M. 1992c. The relationship between gender, comprehension, processing strategies, and cognitive and affective response in foreign-language listening. *Modern Language Journal* 76 (2):160-78.

Bacon, S.M., and M.D. Finnemann. 1990. A study of the attitudes, motives, and strategies of university foreign language students and their disposition to authentic oral and written input. *Modern Language Journal* 74 (4):459-73.

Baddock, B. 1991. Film authenticity and language teaching. *Language Learning Journal* 3:16-8.

Baltova, I. 1994. The impact of video in the comprehension skills of core French students. *Canadian Modern Language Review* 50 (3):507-31.

Beatty, M.J., R.R. Behnke, and L.S. Henderson. 1980. An empirical validation of the Receiver Apprehension Test as a measure of trait listening anxiety. *Western Journal of Speech Communication* 44 (2):132-36.

Berne, J.E. 1995. How does varying pre-listening activities affect second language listening comprehension? *Hispania* 78 (2):316-29.

Bernhardt, E.B., and M.L. Kamil. 1995. Interpreting relationships between L1 and L2 reading: consolidating the linguistic threshold and the linguistic interdependence hypotheses. *Applied Linguistics* 16:15-34.

Berquist, B. 1994. Working memory constraints in L2 listening comprehension. Paper read at European Association for Second Language Acquisition Conference, at Aix-en-Provence, France.

Bialystok, E. 1990. *Communication Strategies: a Psychological Analysis of Second-Language Use*. Oxford: Blackwell.

Black, J.B., and R. Wilensky. 1979. An evaluation of story grammars. *Cognitive Science* 3:213-29.

Blau, E.K. 1990. The effect of syntax, speed and pauses on listening comprehension. *TESOL Quarterly* 24 (4):746-56.

Boyle, E.A., and others. 1994. The effects of visibility on dialogue and performance in a co-operative problem-solving task. *Language and Speech* 37 (1):1-20.

Bransford, J.D., and M.K. Johnson. 1972. Contextual prerequisites for understanding: some investigation of comprehension and recall. *Journal of Verbal Learning and Verbal Behaviour* 11:719-26.

Breen, M.P. 1985. Authenticity in the language classroom. *Applied Linguistics* 6 (1).

Brett, P. 1997. A comparative study of the effects of the use of multimedia on listening comprehension. *System* 25 (1):39-53.

Brindley, G. 1998. Assessing Listening Abilities. *Annual Review of Applied Linguistics* 18:171-191.

Brown, A.L., and A.S. Palincsar. 1982. Inducing strategic learning from texts by means of informed, self-control training. *Topics in Learning and Learning Disabilities* 2 (1):1-17.

Brown, G. 1994. Modes of understanding. In *Language and Understanding*, edited by G. Brown, K. Malkmkjaer, A. Pollitt and J. Williams. Oxford: Oxford University Press.

Brown, G., A.H. Anderson, N. Shadbolt, and A.J. Linch. 1987. Listening comprehension skills in secondary pupils: Scottish Education Department.

Brown, G., and G. Yule. 1983a. *Discourse Analysis, Cambridge Texbooks in Linguistics*. Cambridge: Cambridge University Press.

Brown, G., and G. Yule. 1983b. *Teaching the Spoken Language. An Approach Based on the Analysis of Conversational English*. Cambridge: Cambridge University Press.

Buck, G. 1990. The testing of second language listening comprehension. Unpublished Ph.D. thesis, University of Lancaster.

Buck, G. 1995. How to become a good listening teacher. In *A guide for the teaching of Second Language Listening.*, edited by D. J. Mendelsohn and J. Rubin. San Diego, California: Dominie Press Inc.

Butterworth, B. 1980. Evidence from pauses in speech. In *Language Production*, edited by B. Butterworth. London: Academic Press.

Call, M.E. 1985. Auditory short-term memory, listening comprehension and the input hypothesis. *TESOL Quarterly* 19 (4):765-81.

Cambier, A., and M. Wright. 1995. Satellite television for language learning: how it works, what programmes? What equipment? *Francophonie* 11:21-4.

Campbell, C., and H. Kryszewska. 1992. *Learner-Based Teaching*. Edited by A. Maley, *Resource Books for Teachers*. Oxford: Oxford University Press.

Carrell, P. 1985. Facilitating ESL reading by teaching text structure. *TESOL Quarterly* 19 (4):727-52.

Carrell, P. 1989. Metacognitive awareness and second language reading. *Modern Language Journal* 73:121-34.

Carrell, P. 1992. Awareness of text structure: effects on recall. *Language Learning* 42 (1):1-20.

Carrell, P., L. Devine, and D. Eskey, eds. 1988. *Interactive Approaches to Second Language Reading*. Cambridge: Cambridge University Press.

Cauldwell, R.T. 1996. Direct encounters with fast speech on CD-audio to teach listening. *System* 24 (4):521-28.

Cervantes, R., and G. Gainer. 1992. The effects of syntactic simplification and repetition on listening comprehension. *TESOL Quarterly* 26 (4):767-70.

Chafe, W. 1985. Linguistic differences produced by differences between speaking and writing. In *Literacy, Language and Learning: The Nature and Consequences of Reading and Writing*, edited by D. R. Olson, N. Torrance and A. Hildyard. Cambridge: Cambridge University Press.

Chamot, A.U. 1995. Learning strategies and listening comprehension. In *A Guide for the Teaching of Second Language Listening.*, edited by D. J. Mendelsohn and J. Rubin. San Diego, California: Dominie Press Inc.

Chamot, A.U., and L. Kupper. 1989. Learning strategies in foreign language instruction. *Foreign Language Annals* 22 (1):13-24.

Chaudron, C. 1983. Simplification of input: topic and reinstatements and their effects on L2 learners' recognition and recall. *TESOL Quarterly* 17 (3):437-58.

Chaudron, C., and J.C. Richards. 1986. The effect of discourse markers on the comprehension of lectures. *Applied Linguistics* 7 (2):113-27.

Chavez, M. 1998. Learners' perspectives on authenticity. *IRAL* 36 (4):277-306.

Cheng, P.W. 1985. Restructuring versus automaticity: Alternative accounts of skill acquisition. *Psychological Review* 92:414-43.

Chiang, C.S., and P. Dunkel. 1992. The effect of speech modification, prior knowledge, and listening proficiency on EFL lecture learning. *TESOL Quarterly* 26 (2):345-74.

Chien, C.N. 1998. Listening strategies and processes of Chinese learners of English: a case study of intermediate learners in Taiwan. Unpublished Ph.D. thesis, University of Newcastle upon Tyne, Newcastle upon Tyne.

Clarke, D. 1989. Listening comprehension and the use of authentic materials. *UEA Papers in Linguistics* 29 (30):1-21.

Coniam, D. 2001. The use of audio or video comprehension as an assessment instrument in the certification of Englist language teachers: a case study. *System* 29:1-14.

Conrad, L. 1989. The effects of time-compressed speech on native and EFL listening comprehension. *Studies in Second Language Acquisition* 11 (1):1-16.

Craik, F.I.M., and R.S. Lockhart. 1972. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behaviour* 11:671-84.

Crookes, G. 1990. The utterance and other basic units for second language discourse analysis. *Applied Linguistics* 11 (2):183-99.

Cutler, A. 1994. Segmentation problems, rhythmic solutions. *Lingua* 92:81-104.

Daly, J. 1991. Understanding communication apprehension: an introduction for language educators. In *Language Anxiety: from Theory and Research to Classroom Implications*, edited by E. K. Horwitz and D. Young. Englewood Cliffs, N.J.: Prentice Hall.

de Guerrero, M.C.M., and O.S. Villamil. 2000. Activating the ZPD: Mutual scaffolding in L2 peer revision. *Modern Language Journal* 84 (1):51-68.

Defilippis, D.A. 1980. A study of the Listening Strategies Used by Skillful and Unskillful College French Students in Aural Comprehension Tasks. Unpublished Ph.D. thesis, University of Pittsburgh.

Delabatie, B.D., and D.C. Bradley. 1995. Resolving word boundaries in spoken French. Native and nonnative strategies. *Applied Psycholinguistics* 16 (1):59-81.

Denis, M. 1994. De l'utilisation du document oral authentique en classe de FLE. *Revue de Phonétique Appliquée* 111:91-108.

Dickinson, L. 1970. The language laboratory and advanced teaching. *ELT* 25 (1):32-42.

Dickinson, L. 1987. *Self-Instruction in Language-Learning*. Cambridge: Cambridge University Press.

Dickson, M. 1996. Promoting autonomy through a listening curriculum. In *Promoting Learner Autonomy in University Language Teaching*, edited by E. Broady and M. Kenning. London: CILT.

Drullman, R., and A.W. Bromkhorst. 2000. Multichannel speech intelligibility and talker recognition using monaural, binaural, and three-dimensional auditory presentation. *Journal of the Acoustic Society of America* 107 (4):2224-35.

Dufour, R., and J.F. Kroll. 1995. Matching words to concepts in two languages: A test of the concept mediation model of bilingual representation. *Memory and Cognition* 23:166-80.

Dunkel, P. 1986. Developing listening fluency in L2. Theoretical principles and pedagogical considerations. *Modern Language Journal* 70 (2):99-106.

Dupuy, B.C. 1999. Narrow listening: an alternative way to develop and enhance comprehension in students of French as a foreign language. *System* 27:351-61.

Ellis, G., and B. Sinclair. 1989. *Learning to Learn English. A Course in Learner Training*. Cambridge: Cambridge University Press.

Ely, P. 1984. *Bring the Lab back to Life*. Oxford: Pergamon Press.

Ericsson, K.A., and H.A. Simon. 1980. Verbal reports as data. *Psychological Review* 87:215-51.

Ericsson, K.A., and H.A. Simon. 1984. *Protocol Analysis: Verbal Reports as Data*. Cambridge, MA: MIT Press.

Ericsson, K.A., and H.A. Simon. 1987. Verbal reports on thinking. In *Introspection in Second Language Research*, edited by C. Faerch and G. Kasper. Clevedon: Multilingual Matters.

Fernández-Toro, M. 1999. *Training Learners for Self-Instruction*. London: CILT.

Fernández-Toro, M., and F.R. Jones. 2001. *D.I.Y. Techniques for Language Learners*. London: CILT.

Field, J. 1997. Notes on listening: Variability and assimilation. *Modern English Teacher* 6 (1):51-2.

Field, J. 1998. Skills and strategies: towards a new methodology for listening. *ELT Journal* 52 (2):110-18.

Fisher, B., D. Lynch, and C. Allen. 1995. *Satellite Television in the Classroom*. London: CILT.

Forster, K.I. 1976. Accessing the mental lexicon. In *New Approaches to Language Mechanisms*, edited by R. J. Wales and E. Walker. Amsterdam: North-Holland.

Forster, K.J. 1979. Levels of processing and the structure of the language processor. In *Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett*, edited by W. E. Cooper and E. C. T. Walker. Hillsdale, N.J.: Erlbaum.

Foulke, E., and T.G. Sticht. 1969. Review of research on the intelligibility and comprehension of accelerated speech. *Psychological Bulletin* 72:50-62.

Garnham, A. 1994. Psychological processes and understanding. In *Language and Understanding*, edited by G. Brown, K. Malkmkjaer, A. Pollitt and J. Williams. Oxford: Oxford University Press.

Geddes, M., and R. White. 1978. The use of semi-scripted simulated authentic speech and listening comprehension. *Audio-Visual Language Journal* 16 (3):134-45.

Gilbert, J. 1995. Pronunciation practice as an aid to listening comprehension. In *A Guide for the Teaching of Second Language Listening.*, edited by D. J. Mendelsohn and J. Rubin: Dominie Press Inc.

Goh, C. 1997. Metacognitive awareness and second language listeners. *ELT Journal* 51 (4):361-69.

Goh, C.C.M. 1998. How ESL Learners with Different Listening Abilities Use Comprehension Strategies and Tactics. *Language Teaching Research* 2 (2):124-47.

Grabe, W. 1991. Current developments in second language reading research. *TESOL Quarterly* 25 (3):375-406.

Greene, J. 1987. *Memory, Thinking and Language*. London & New York: Methuen.

Gregory, M. 1967. Aspects of varieties differentiation. *Journal of Linguistics* 3 (2):177-274.

Griffiths, R. 1990. Speech rate and non-native speaker comprehension: a preliminary study in time-benefit analysis. *Language Learning* 40 (3):311-36.

Griffiths, R. 1991. Pausological research in an L2 context: a rationale and review of selected studies. *Applied Linguistics* 12:345-64.

Halliday, M.A.K. 1989. *Spoken and Written Language*. Oxford: Oxford University Press.

Halliday, M.A.K., and R. Hasan. 1985. *Language, Context and Text: Aspects of Language in a Social-Semiotic Perspective.* Oxford: Oxford University Press.

Harrington, M., and M. Sawyer. 1992. L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition* 14 (1):25-38.

Hart, I. 1992. Video, foreign language teaching and the documentary tradition. *System* 20 (1):1-13.

Hayashi, T. 1991. Interactive processing of words in connected speech in L1 and L2. *IRAL* 29 (2):151-60.

Henrichsen, L.E. 1984. Sandhi-variation: a filter of input for learners of ESL. *Language Learning* 34 (3):103-26.

Herron, C., M. Morris, T. Secules, and L. Curtis. 1995. A comparison study of the effects of video-based versus text-based instruction in the foreign language classroom. *French Review* 68 (5):775-795.

Herron, C., and I. Seay. 1991. The effect of authentic oral texts on student listening comprehension in the foreign language classroom. *Foreign Language Annals* 24 (6):487-91.

Hildyard, A., and D. Olson. 1978. Memory and inference in the comprehension of oral and written discourse. *Discourse Processes* 1:91-117.

Hildyard, A., and D. Olson. 1982. On the comprehension and memory of oral vs. written discourse. *Advances in Discourse Processes* 9:19-34.

Hill, B. 1991. *Making the Most of Satellites and Interactive Video.* London: CILT.

Horiba, Y., V.d. Broek, and Fletcher. 1993. Second language readers' memory for narrative texts: evidence for structure-preserving top-down processing. *Language Learning* 43 (3):345-72.

Horwitz, E.K., M.B. Horwitz, and J.A. Cope. 1991. Foreign language classroom anxiety. In *Language Anxiety: from Theory and Research to Classroom Implications*, edited by E. K. Horwitz and D. Young. Englewood Cliffs, N.J.: Prentice Hall.

Horwitz, E.K., and D. Young, eds. 1991. *Language Anxiety: from Theory and Research to Classroom Implications.* Englewood Cliffs, N.J.: Prentice Hall.

Hossein, N., and A. Cumming. 2000. What's in a ZPD? A case study of a young ESL student and teacher interacting through dialogue journals. *Language Teaching Research* 4 (2):95-121.

Institute of Linguists Educational Trust. 2000. *Examinations in Languages for International Communication (ELIC).* London: Institute of Linguists Educational Trust.

Janssen, J., and H. van Loon. 1991. *New Media in the Humanities.* Amsterdam: University of Amsterdam.

Jones, T.J.R. 1962. Testing oral proficiency by tape. *Education* 21:407-09.

Just, M., and P. Carpenter. 1992. A capacity hypothesis of comprehension: individual differences in working memory. *Psychological Review* 99 (122):149.

Just, M.A., and P.A. Carpenter. 1977. *Cognitive Processes in Comprehension.* New Jersey: Erlbaum Associates.

Just, M.A., and P.A. Carpenter. 1980. A theory of reading: From eye fixations to comprehension. *Psychological Review* 87:329-54.

Kasper, G., and E. Kellerman, eds. 1997. *Communication Strategies. Psycholinguistic and Sociolinguistic Perspectives.* Edited by G. Kasper and E. Kellerman. London: Longman.

Kasper, G.K.E. 1997. Introduction: approaches to communication strategies. In *Communication Strategies*, edited by G. Kasper and E. Kellerman. London: Longman.

Kellerman, S. 1990. Lip service: the contribution of the visual modality to speech perception. *Applied Linguistics* 11 (3):272-80.

Kellermann, S. 1992. 'I see what you mean': the role of kinesic behaviour in listening and implications for foreign second language learning. *Applied Linguistics* 13 (3):239-58.

Kelly, P. 1991. Lexical ignorance: the main obstacle to listening comprehension with advanced foreign language learners. *IRAL* 29 (2):135-49.

Kess, J.F. 1992. *Psycholinguistics. Psychology, Liguistics, and the Study of Natural Language.* Edited by E. F. Konrad Koerner. Vol. 86, *Series IV - Current Issues in Linguistic Theory.*

Klein-Braley, C. 1985. A cloze-up on the C-test: A study in the construct validation of authentic tests. *Language Testing* 2:76-104.

Krashen, S. 1981. The case for narrow reading. *TESOL Newsletter* 12:23.

Krashen, S. 1985. *The Input Hypothesis. Issues and Implications.* London: Longman.

Krashen, S. 1996. The case for narrow listening. *System* 24 (1):97-100.

Kroll, B. 1977. Combining ideas in written and spoken English: a look at subordination and coordination. *Southern California Occasional Papers in Linguistics* 5 (Discourse across Time and Space):69-108.

Lancien, T. 1986. *Le Document Vidéo, Techniques de Classe.* Paris: Clé International.

Laviosa, F. 1990. A preliminary investigation of the listening problem-solution process and the strategies of five advanced learners of Italian as a second language. *Journal for Language Teaching* 25 (3):17-29.

Laviosa, F. 1991a. An investigation of the listening strategies of advanced learners of Italian as a second language. Paper read at Bridging Theory and Practice in the Foreign Language Classroom, October 1991, at Maryland.

Laviosa, F. 1991b. A Preliminary Investigation of the Listening Problem-Solution Process and the Strategies of Five Advanced Learners of Italian as a Second Language. Ph.D. thesis, State University of New York at Buffalo.

Laviosa, F. 1993. Second language listening strategy training. *Italiana* 5:295-308.

Levy, E.T., and D. McNeill. 1992. Speech, gesture and discourse. *Discourse Processes* 15 (3):277-301.

Lewkowicz, J.A. 2000. Authenticity in language testing: some outstanding questions. *Language Testing* 17 (1):43-64.

Littlewood, W. 1981. *Communicative Language Teaching: an Introduction.* Cambridge: Cambridge University Press.

London Chamber of Commerce and Industry. 1999. *Foreign Languages for Industry and Commerce (FLIC). Extended Syllabus.* London: London Chamber of Commerce and Industry.

Lonergan, J. 1982. Video recordings in language teaching. *Media in Education and Development* 15 (1).

Lonergan, J. 1983. The value of video activities for secondary school language teachers. *Practical English Teacher* (June).

Lonergan, J. 1984. *Video in Language Teaching, New Directions in Language Teaching.* Cambridge: Cambridge University Press.

Long, D.R. 1989. Second language listening comprehension: a schema theoretic perspective. *Modern Language Journal* 73 (1):32-39.

Long, D.R. 1991. Listening processes and authentic texts. Southern Conference on Language Teaching. ERIC Report ED348851.

Long, M.H. 1983. Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics* 4:126-41.

Lund, R.J. 1991a. A comparison of second language reading and listening comprehension. *Modern Language Journal* 75 (2):196-204.

Lund, R.J. 1991b. The effects of listening tasks on comprehension. Paper read at Research Perspectives in Adult Language Learning and Acquisition Annual Symposium, October 11-12, at Columbus, OH.

Lynch, A.J. 1989. Grading foreign language listening comprehension materials: the use of naturally modified interaction. Ph.D. thesis, University of Edinburgh.

Lynch, T. 1995. The development of interactive listening strategies in second language academic settings. In *A Guide for the Teaching of Second Language Listening*, edited by D. J. Mendelsohn and J. Rubin. San Diego, California: Dominie Press Inc.

Lynch, T. 1998. Theoretical perspectives on listening. *Annual Review of Applied Linguistics* 18:3-19.

Mandler, J.M. 1982. Recent research on story grammars. In *Language and Comprehension*, edited by J. F. L. Ny and W. Kintsch. Amsterdam: North-Holland.

Mandler, J.M., and N.S. Johnson. 1977. Remembrance of things parsed: story structure and recall. *Cognitive Psychology* 9:111-51.

Marslen-Wilson, W., and L.K. Tyler. 1980. The temporal structure of spoken language understanding. *Cognition* 8:1-71.

Marslen-Wilson, W., and A. Welsh. 1978. Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology* 10:29-63.

Matsumoto, K. 1993. Verbal-report data and introspective methods in second language research: state of the art. *RELC Journal* 24 (1):32-60.

Matsumoto, K. 1994. Introspection, verbal reports and second-language learning strategy research. *Canadian Modern Language Review* 50 (2):363-86.

Matter, J.F. 1989. The language laboratory revisited. Paper read at Man and the Media Symposium, at Frankfurt.

McCarthy, M. 1991. *Discourse Analysis for Language Teachers.* Cambridge: Cambridge University Press.

McClelland, J., and J.L. Elman. 1986. Interactive processes in speech perception: The TRACE model. In *Parallel Distributed Processing: Explorations in the*

Microstructure of Cognition, edited by J. McClelland, D. E. Rumelhart and the PDP Research Group. Cambridge, MA: MIT Press.

McClelland, J.L., and D.E. Rumelhart. 1981. An interactive-activation model of context effects in letter perception. Part I: An account of basic findings. *Psychological Review* 88:375-407.

McClelland, J.L., D.E. Rumelhart, and G.E. Hinton. 1986. The appeal of parallel distributed processing. In *Parallel Distributed Processing. Explorations in the Microstructure of Cognition*, edited by D. E. Rumelhart, J. L. McClelland and the PDP Research Group. Cambridge, Mass.: MIT Press.

McDonough, S.H. 1999. Learner strategies - State of the art article. *Language Teaching* 32:1-18.

McGovern, J. 1984. Video Applications in ELT. *ELT Documents* 114:57-67.

McGregor, G. 1986. Listening outside the participation framework. In *The Art of Listening*, edited by G. McGregor and R. S. White. London: Croon Helm.

McLaughin, B. 1990. Restructuring. *Applied Linguistics* 11 (2):113-28.

Meinhof, U.H., and M. Bergman. 1991a. Interactive video as a tool for text comprehension. An experiment with television news. In *New Media in the Humanities*, edited by Janssen and van Loon. Amsterdam: University of Amsterdam.

Meinhof, U.H., and M. Bergman. 1991b. *ITN World News*. Oxford: Oxford University Press.

Mendelsohn, D. 1995. Applying learning strategies in the second/foreign language listening comprehension lesson. In *A Guide for the Teaching of Second Language Listening*, edited by D. J. Mendelsohn and J. Rubin. San Diego, California: Dominie Press Inc.

Mendelsohn, D.J. 1998. Teaching Listening. *Annual Review of Applied Linguistics* 18:81-101.

Merlet, S., and D. Gaonac'h. 1995. Mise en évidence de stratégies compensatoires dans la compréhension orale en L.E. *Revue de Phonétique Appliquée* 115 (17):273-92.

Meyer, B.F.J. 1975. *The Organisation of Prose and its Effects on Memory*. Amsterdam: North-Holland.

Miller, G.A. 1956. The magical number seven, plus or minus two: Some limits in our capacity for processing information. *Psychological Review* 63:81-97.

Minsky, M. 1977. Frame-system theory. In *Thinking: Readings in Cognitive Science*, edited by P. N. Johnson-Laird and P. C. Wason. Cambridge: Cambridge University Press.

Morton, J. 1969. Interaction of information in word recognition. *Psychological Review* 76:165-78.

Morton, J. 1979. Word Recognition. In *psycholinguistics 2: Structures and Processes*, edited by J. Morton and D. J. Marshall. Cambridge, Mas.: MIT Press.

Morton, J. 1982. Disintegrating the lexicon: An information processing approach. In *On Mental Representation*, edited by W. W. J. Mehler and M. Garrett. Hillsdale, N.J.: Erlbaum.

Mueller, T. 1959. Auding tests. *Modern Language Journal* 43:185-7.

Nisbett, R.E., and T.D. Wilson. 1977. Telling more than we can know: Verbal reports on mental processes. *Psychological Review* 84:231-59.

Norris, D., J. McQueen, and A. Cutler. 1995. Competition and segmentation in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition* 21:1209-28.

Nuttall, C. 1996. *Teaching Reading Skills in a Foreign Language*. London: Heinemann.

O'Malley, J.M., A.U.Chamot, G. Stewner-Manzanares, B.P. Russo and. L. Küpper 1985a. Learning strategies used by beginning and intermediate ESL students. *Language Learning* 35:21-46.

O'Malley, J.M., A. Chamot, and L. Küpper. 1989. Listening comprehension strategies in second language acquisition. *Applied Linguistics* 10 (4):418-37.

O'Malley, J.M., and A.U. Chamot. 1990. *Learning Strategies in Second Language Acquisition*. Cambridge: Cambridge University Press.

O'Malley, J.M., A.U. Chamot, G. Stewner-Manzanares, B.P. Russo, and L. Küpper. 1985b. Learning strategy applications with students of English as a second language. *TESOL Quarterly* 19:285-96.

Oxford Cambridge and RSA Examinations (OCR). 2000a. *Advanced Subsidiary GCE. Specimen Materials, Spanish. Listening, Reading and Writing 1*. Cambridge: OCR.

Oxford Cambridge and RSA Examinations (OCR). 2000b. *Advanced Subsidiary GCE. Specimen Materials, Spanish. Listening, Reading and Writing 2*. Cambridge: OCR.

Oxford Cambridge and RSA Examinations (OCR). 2000c. *General Certificate of Secondary Education. Specimen Materials, Spanish*. Cambridge: OCR.

Oxford, R., Y. Park-Oh, S. Ito, and M. Sumarall. 1993. Learning a language by satellite television: what influences student achievement? *System* 21 (1):31-48.

Oxford, R.L. 1990. *Language Learning Strategies. What every Teacher should Know*. Boston, Mass.: Heinle & Heinle publishers.

Parry, T. 1984. Videotape vs. audiotape for listening comprehension tests: an experiment. *OMLTA Journal*:47-53.

Peterson, P.W. 1985. The bridge course: listening comprehension in authentic settings. *TESOL Newsletter* 19 (6):21.

Pisoni, D.B., and P.A. Luce. 1987. Acoustic-phonetic representation in word recognition. In *Spoken Word Recognition*, edited by U. H. Frauenfelder and L. K. Tyler. Cambridge, Mass.: MIT Press.

Porter, D., and J. Roberts. 1981. Authentic listening activities. *ELT Journal* 36 (1):37-47.

Rabinowitz, M., and M.T. Chi. 1987. An interactive model of strategic processing. In *Handbook of Cognitive, Social, and Neuropsychological Aspects of Learning Disabilities*, edited by S. J. Ceci. Hillsdale, N.J.: Erlbaum.

Reed, S.K. 1991. *Cognition*. Pacific Grove, California: Brooks/Cole Publishing Company.

Rees-Miller, J. 1993. A critical appraisal of learner-training: theoretical bases and teaching implications. *TESOL Quarterly* 27 (4):679-89.

Rings, L. 1986. Authentic language and authentic conversational texts. *Foreign Language Annals* 19 (3):203-8.

Rivenc, M.M. 1995. La compréhension orale chez les débutants: stratégies de réception, stratégies d'enseignement. *Revue de Phonétique Appliquée* 115 (17):331-43.

Rixon, S. 1986. *Developing Listening Skills*. London: Macmillam.

Rose, R.G. 1995. French language satellite TV in the classroom. *Foreign Language Annals* 28 (4):518-26.

Ross, S. 1997. An introspective analysis of listener inferencing on a second language listening task. In *Communication Strategies*, edited by G. Kasper and E. Kellerman: Longman.

Rost, M. 1990. *Listening in Language Learning*. London: Longman.

Rost, M., and S. Ross. 1991. Learner use of strategies in interaction: typology and teachability. *Language Learning* 41 (2):235-73.

Rowntree, D. 1986. *Teaching through Self-Instruction*. London & New York: Kogan Page/Nichols Publishing.

Rubin, J. 1975. What the "good language learner" can teach us. *TESOL Quarterly* 9 (1):41-51.

Rubin, J. 1981. Study of cognitive processes in second language learning. *Applied Linguistics* 11:117-131.

Rubin, J. 1990. Improving foreign language listening comprehension. In *Georgetown University Round Table*, edited by J. E. Alatis. Washington, DC: Georgetown University Press.

Rubin, J. 1995. The contribution of video to the development of competence in listening. In *A guide for the Teaching of Second Language Listening*, edited by D. J. Mendelsohn and J. Rubin. San Diego, California: Dominie Press Inc.

Rubin, J., J. Quinn, and J. Enos. 1988. Improving foreign language listening comprehension. Washington, DC: International Research and Studies Program. U.S. Department of Education.

Rubin, J., and I. Thompson. 1992. Materials selection in strategy instruction for Russian listening comprehension: George Washington University.

Rubin, J., and I. Thompson. 1994. *How to Be a More Successful Language Learner*. Boston, Mass.: Heinle & Heinle Publishers.

Rumelhart, D.E. 1981. Schemata: the building blocks of cognition. In *Comprehension and Teaching: Research Reviews*, edited by J. T. Guthrie. Newark: International Reading Association.

Rushoff, B., and D. Wolff. 1991. Developing and using interactive audio for foreign language learning. In *New Media in the Humanities*, edited by J. Janssen and H. van.Loon. Amsterdam: University of Amsterdam.

Sanford, A.J., and S.C. Garrod. 1981. *Understanding Written Language*. Chichester: Wiley.

Schank, R.C., and R.P. Abelson. 1977a. Scripts, plans and knowledge. In *Thinking: Readings in Cognitive Science*, edited by P. N. Johnson-Laird and P. C. Wason. Cambridge: Cambridge University Press.

Schank, R.C., and R.P. Abelson. 1977b. *Scripts, Plans, Goals and Understanding.* Hillsdale, N.J.: Lawrence Erlbaum.

Schmidt, R. 1990. The role of consciousness in second language learning. *Applied Linguistics* 11:129-158.

Schmidt, R. 1992. Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition* 14:57-85.

Schmidt, R. 1994. Deconstructing consciousness in search of useful definitions for applied linguistics. In *Consciousness in Second Language Learning*, edited by J. H. Hulstijn and R. Schmidt: International Association of Applied Linguistics (AILA).

Schneider, W., and R.M. Shiffrin. 1977. Controlled and automatic human information processing: I. Detection, search and attention. *Psychological Review* 84 (1):1-57.

Schrafnagl, J., and D. Cameron. 1988. Are you decoding me? The assessment of understanding in oral interaction. In *Applied Linguistics in Society*, edited by P. Grunwell. London: CILT.

Seliger, H. 1977. Does practice make perfect? A study of interaction patterns and second language competence. *Language Learning* 27 (2):263-78.

Seliger, H.W. 1983. The language learner as linguist: Of metaphors and realities. *Applied Linguistics* 4:179-91.

Sherman, J. 1997. The effect of question preview in listening comprehension tests. *Language Testing* 14 (2):185-213.

Shiffrin, R.M., and W. Schneider. 1977. Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review* 84 (2):127-90.

Shohamy, E., and O. Inbar. 1988. Content and construct validity of listening comprehension tests: the effect of text and question type. Paper read at 10th Annual Language Testing Research Colloquium, March 1988, at Urbana, Illinois.

Shohamy, E., and O. Inbar. 1991. Validation of listening comprehension tests. *Language Testing* 8 (1):23-40.

Spencer, R.E. 1964. The relative effectiveness of earphones and loudspeakers as a means of presenting a listening test in a foreign language. *Modern Language Journal* 48:346-49.

Spielberger, C.D., R.L. Gorush, and R.E. Lushene. 1968. *The State-Trait Anxiety Inventory (STAI).* Tallahassee, Florida: Florida State University.

Stempleski, S., and B. Tomalin. 1990. *Video in Action. Recipes for Using Video in Language Teaching.* Englewood Cliffs, N.J.: Prentice Hall.

Stillman, R. 1980. Auditory brain mechanisms. In *Auditory Processing and Language: Clinical and Research Prespective*, edited by P. Levinson and C. Sloan. New York & London: Grune & Stratton.

Strange, J., and D. Strange. 1991. Survey: video materials. *ELT Journal* 45 (4):335-59.

Tannen, D. 1980. Implications of the oral/literate continuum for cross-cultural communication. In *Currrent Issues in Bilingualism (Georgetown University Round Table on Languages and Linguistics, 1980)*, edited by J. Alatis. Washington, D.C.: Georgetown University Press.

Tannen, D. 1982. The oral/literate continuum in discourse. In *Spoken and Written Language: Exploring Orality and Literacy*, edited by D. Tannen. Norwood, NJ: Ablex.

Tannen, D. 1985. Relative focus of involvement in oral and written discourse. In *Literacy, Language and Learning*, edited by D. R. Olson, N. Torrance and A. Hildyard. Cambridge: Cambridge University Press.

Tarone, E. 1981. Some thoughts on the notion of communication strategy. *TESOL Quarterly* 15:285-95.

Tauroza, S., and D. Allison. 1994. Expectations-driven understanding in information systems lecture comprehension. In *Academic Listening - Research Perspectives*, edited by J. Flowerdew. Cambridge: Cambridge University Press.

Teichert, H.U. 1996. A comparative study using illustrations, brainstorming, and questions as advance organisers in intermediate college German conversation classes. *Modern Language Journal* 80 (4):509-17.

Thompson, I. 1995. Assessment in second/foreign language listening comprehension. In *A guide for the teaching of Second Language Listening.*, edited by D. J. Mendelsohn and J. Rubin. San Diego, California: Dominie Press Inc.

Thompson, I., and J. Rubin. 1993. Improving Listening Comprehension in Russian: Report submitted to US Dept. of Education, International Research and Studies Program. Grant #PO17A00032.

Thompson, I., and J. Rubin. 1996. Can strategy instruction improve listening comprehension? *Foreign Language Annals* 29 (3):331-42.

Thorndyke, P.W. 1977. Cognitive studies in comprehension and memory of narrative discourse. *Cognitive Psychology* 9:77-110.

Trabasso, T. 1981. On the making of inferences during reading comprehension and their assessment. In *Comprehension and Teaching*, edited by J.T. Guthrie: Newark, Delaware: International Reading Association.

Trabasso, T., and J.P. Magliano. 1996. Conscious understanding during comprehension. *Discourse Processes* 21:255-87.

Turner, K. 1995. *Listening in a Foreign Language. A skill we take for granted?* Vol. 26, *Pathfinder.* London: CILT.

Tyler, A. 1994. The role of repetition in perceptions of discourse coherence. *Journal of Pragmatics* 21 (6):671-88.

Tyler, L.K., and U.H. Frauenfelder. 1987. The process of word recognition: an introduction. In *Spoken Word Recognition*, edited by L. Tyler and K. Frauenfelder. Cambridge, Mass.: MIT Press.

Underwood, M. 1989. *Teaching Listening.* London: Longman.

Ur, P. 1984. *Teaching Listening Comprehension.* Cambridge: Cambridge University Press.

Vandergrift, L. 1996. Listening comprehension strategies of core French high-school students. *Canadian Modern Language Review* 52 (2):200-23.

Vandergrift, L. 1998. Successful and less successful listeners in French: What are the strategy differences? *French Review* 71 (3):370-395.

Vanderplank, R. 1993. 'Pacing' and 'spacing' as predictors of difficulty in speaking and understanding English. *ELT Journal* 47 (2):117-25.

Vann, R.J., and R.G. Abraham. 1990. Strategies of unsuccessful language learners. *TESOL Quarterly* 24 (2):177-98.

Vihman, M. 1997. The infant listener. Paper read at IATEFL ReSIG Listening Skills conference, at Cambridge.

Vogely, A. 1995. Perceived strategy use during performance on three authentic listening comprehension tasks. *Modern Language Journal* 79 (1):41-56.

Vogely, A.J. 1998. Listening comprehension anxiety: Students' reported sources and solutions. *Foreign Language Annals* 31 (1):67-80.

Voss, B. 1979. Hesitation phenomena as sources of perceptual errors for non-native speakers. *Language and Speech* 22:129-44.

Vygotsky, L.S. 1978. Interaction between learning and development. In *Mind in Society. The Development of Higher Psychological Processes*, edited by M. Cole, V. John-Steiner, S. Schribner and E. Souberman. Cambridge, Massachusetts: Harward University Press.

Wardrop, E., and K. Anderson. 1992. An investigation into learner's reactions to learner training for self-access listening. *Edinburgh Working Papers in Applied Linguistics* 3:134-45.

Weinstein, C.E., and R.E. Mayer. 1986. The teaching of learning strategies. In *Handbook on Research on Teaching*, edited by M. R. Wittrock. New York: Macmillan.

Weissenrieder, M. 1987. Listening to the news in Spanish. *Modern Language Journal* 71 (1):18-27.

Wenden, A. 1995. Learner training in context: a knowledge-based approach. *System* 23 (2):183-94.

Wheeless, L.R. 1975. An investigation of receiver apprehension and social context dimensions of communication apprehension. *Speech Teacher* 24:261.

Widdowson, H. 1979. *Language Purpose and Language Use*. Oxford: Oxford University Press.

Widdowson, H.G. 1984. *Explorations in Applied Linguistics*. 2 vols. Vol. 2. Oxford: Oxford University Press.

Willis, J. 1983. The role of the visual element in spoken discourse: implications for the exploitation of video in the EFL classroom. *ELT Documents* 114 (Video Applications in English Language Teaching):29-42.

Wright, T. 1987. *Roles of Teachers and Learners*. Oxford: Oxford University Press.

Zhao, Y. 1997. The effects of listener's control of speech rate on second language comprehension. *Applied Linguistics* 18 (8):49-68.

Zwaan, R., and C. Brown. 1996. The influence of language proficiency and comprehension skill on situation-model construction. *Discourse Processes* 21:289-327.

Zwaan, R.A., M.C. Langston, and A.C. Graesser. 1995. The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science* 6:292-7.