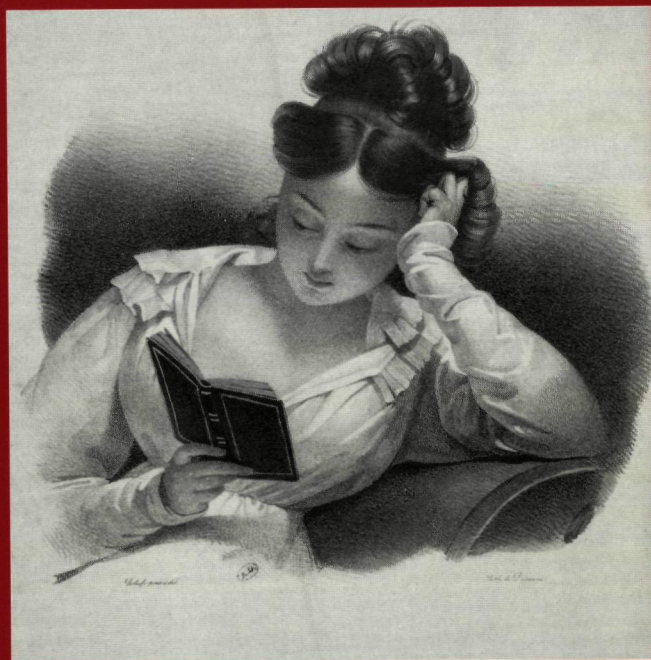


Corine van den Brandt

Context Use and Learning to Read in a Foreign Language



CENTER FOR LANGUAGE STUDIES



Netherlands
Graduate
School of
Linguistics

Landelijke Onderzoekschool Taalwetenschap

Context Use and Learning to Read in a Foreign Language

Context use and learning to read in a foreign language

Een wetenschappelijke proeve op het gebied van de Letteren

Proefschrift

ter verkrijging van de graad van doctor
aan de Katholieke Universiteit Nijmegen,
volgens besluit van het College van Decanen
in het openbaar te verdedigen
op woensdag 22 september 1999,
des namiddags om 3.30 uur precies

door

Jacoba Joanna Maria van den Brandt

geboren op 11 augustus 1962

te 's-Hertogenbosch

Promotor: Prof. Dr. T.J.M. van Els

Co-promotor: Dr. C.J.M. van Esch

Manuscriptcommissie: Prof. Dr. C.A.J. Aarnoutse

Prof. Dr. C.L.J. de Bot

Prof. Dr. M.I.J.M. Goethals (Katholieke
Universiteit Leuven)

Published by

Holland Academic Graphics

P.O. Box 53292

2505 AG The Hague

The Netherlands

phone: +31 70 4480203

fax: +31 70 4480177

e-mail: info@hagpub.com

[http: www.hagpub.com](http://www.hagpub.com)

ISBN 90-5569-110-0

NUGI 941

Copyright © 1999 by Corine van den Brandt. All rights reserved.

Printed in The Netherlands.

Acknowledgements

It goes without saying that a study such as this could not have been undertaken without a great deal of support. I would like to thank the following persons:

Toon van Bommel for his help and enthusiasm during the execution of tests and programmes in Germany.

School boards, staff and students at the 'Berufsbildende Schule' (Vocational Education) in Kleve and Geldern (Germany) for their time and efforts.

Staff of CITO (Dutch National Institute for Educational Measurement), particularly Jan Mets, Margriet Welling and Noud van Zuylen, for their useful advice and permission to use their test materials.

Marisol Hofmans for transcribing the teachers' interviews.

Julie Beardsell and Hilde Nicolassen for checking my English.

The late Eric Schils for his help with the statistics. His creative way of thinking and great sense of humour have been very stimulating.

Núria Domínguez, Andreu van Hooft, Rob le Pair, Agnes Smorenburg and my other colleagues at the Department of Spanish Language and Literature, the Department of Business Communication Studies and the Department of Applied Linguistics.

And last but not least, a special word of thanks goes to Peter van den Kroonenberg, my family and friends.

Contents

Acknowledgements		v
Introduction		1
Chapter 1	Teaching foreign language reading	3
1.1	Importance of foreign language reading	3
1.2	Reading and foreign language teaching	5
1.3	The Netherlands: Teaching practice	7
1.4	The Netherlands: Recent reforms	8
1.5	Summary and conclusion	12
Chapter 2	Reading	13
2.1	Three models of reading	13
2.2	Aspects of interactive reading	15
2.2.1	Knowledge-based processing	16
2.2.2	Text-based processing	20
2.2.3	Compensatory processing	23
2.3	Summary and conclusion	24
Chapter 3	Non-native reading and the use of context	25
3.1	Non-native reading: Review of research results	25
3.1.1	Knowledge-based processing and non-native reading	25
3.1.2	Text-based processing and non-native reading	27
3.1.3	Language threshold hypothesis	28
3.1.4	Conclusions research review	29
3.2	Variables of non-native reading	30
3.2.1	Linguistic knowledge	30
3.2.2	Prior content knowledge	31
3.2.3	Strategy use	33
3.2.4	Metacognition	34
3.2.5	Other variables	35
3.3	A model of non-native reading	36
3.4	Context use	37
3.5	Summary and conclusion	40
Chapter 4	Research plan and execution	43
4.1	The 1987-project	44
4.2	Aim of the follow-up investigation	45
4.3	Pilot experiment	48

CONTENTS

4.4	Research design	49
4.5	Subjects	51
4.6	Planning of the experiment	52
4.7	Experimental condition	53
4.7.1	Spanish training programme	53
4.7.2	English training programme	55
4.7.3	Justification of the contents of the training programmes	56
4.8	Control condition	58
4.9	Instruments	59
4.9.1	Vocabulary tests	59
4.9.2	Context tests	61
4.9.3	Reading comprehension tests	62
4.9.4	Questionnaires	67
4.10	Execution of the main experiment	68
Chapter 5	Research results	71
5.1	Results of the context and reading comprehension tests	71
5.1.1	Overall results	72
5.1.2	Results by school class	75
5.1.3	Results at the individual level	79
5.2	Relevance of four reader-dependent variables	82
5.2.1	Vocabulary	83
5.2.2	Attitude, self-confidence and estimate of context use	84
5.3	Summary and conclusion	85
Chapter 6	Evaluation of programmes and tests	87
6.1	The students' judgement of programmes and tests	87
6.2	The teachers' judgement of programmes and tests	90
6.2.1	The training and control programmes	91
6.2.2	Instruments	95
6.3	Summary and conclusion	97
Chapter 7	Discussion	101
7.1	Summary of results	101
7.2	Discussion of results	107
7.2.1	Training context use	108
7.2.2	Reader-dependent variables	117
7.3	Implications for research on and teaching of non-native reading	118
7.4	Suggestions for further research	122

CONTENTS

References	125
List of Appendices	135
Samenvatting	219
Curriculum Vitae	223

Introduction

This research is concerned with the teaching of foreign language reading. Over the last 15 years or so, there have been considerable developments in this field, with significant contributions from both theorists and practitioners. One of the aspects that has been discussed extensively includes the teaching of reading strategies. Therefore, there is a need for clarity and further research into the effectiveness of existing methodologies and the extent to which newly developed ideas about the teaching of reading have proved to be successful in practice.

The aim of this study is to assess the effects of a training programme in the use of context. The training programme involves two different kinds of contextual knowledge: knowledge of the general context, that is, prior knowledge about the topic of a text and knowledge of the local context of the word in the sentence. The training programme in the use of context focuses on improving reading skills by making learners aware of how they can increase their understanding of a foreign language text by making inferences based on both their background knowledge and the specific textual information itself.

This study is a follow-up to earlier research into the effects of context-use training, carried out with a group of beginners in Spanish (Van Esch, 1987). The results of this earlier research showed positive effects of context-use training.

This follow-up is unique in that, unlike the earlier study which focused on only one foreign language, it included two groups of students, one of which studied Spanish and the other English. The principal advantage of involving two such groups was that it provided the opportunity to assess differences in the effects of a context-use training programme between students of varying proficiency levels: the students learning Spanish were beginners and those learning English were of an intermediate level.

In view of the ecological validity the research was conducted within a classroom environment, just as the earlier project. The effects of context-use training were explored in the settings of the Spanish and English lessons of the same school that was involved in the 1987-project. For control purposes, also a second school was included.

This book is divided into two parts. The first part, comprising chapters one to three inclusive, explores theoretical issues in the field of reading comprehension, in particular with regard to the use of context in the reading of a foreign language. Chapter 1 reviews current ideas about the teaching of foreign language reading. Besides discussing general aspects of the teaching of reading, reference is made to the recent reforms in the curriculum of secondary education in the Netherlands.

Chapter 2 considers models of human information processing and reading. In Chapter 3 some relevant aspects of non-native reading are discussed, the different types of context are distinguished and an explanation is given as to why we believe that context-use training is a relevant factor in the teaching of foreign language reading comprehension.

The second part, comprising chapters four to seven inclusive, focuses on the design, procedures and results of two seven-week training programmes in the use of context: a training programme for learners of Spanish at an elementary level and a training programme for intermediate learners of English. In Chapter 4 the research plan of our field-experiment is outlined and three research questions are formulated. In Chapter 5 the results are presented. It is established whether the context-use training programme for beginning learners of Spanish was successful (research question 1), whether the training programme for intermediate learners of English led to other results than the training programme for beginning learners of Spanish (research question 2) and to what degree the scores on a number of four reader-dependent variables correlated with the performance on context use and reading comprehension (research question 3). These variables are vocabulary knowledge, attitude towards the foreign language, self-confidence and the student's own estimate of the degree to which he or she uses contextual information while reading Spanish or English. Chapter 6 considers the students' and teachers' evaluation of the project. Finally, in Chapter 7, a summary of the findings is presented, the limitations and implications of the results are considered and some suggestions for further research are given.

Chapter 1

Teaching foreign language reading

Except for the evaluation of a training programme in the use of context, one of the goals of this study is to contribute to the development of teaching foreign language reading in general and in secondary education in the Netherlands in particular. Therefore, reference is repeatedly made to foreign language teaching specifically in a Dutch context.

Section 1.1 is concerned with the importance of foreign language reading proficiency. Section 1.2 explores the role of reading within the framework of three influential approaches towards foreign language teaching. In section 1.3 and 1.4 attention is paid to the teaching of foreign language reading in secondary education in the Netherlands. Section 1.3 is about teaching practice and section 1.4 deals with the recent reforms in the curriculum. Section 1.5 is a summary of this chapter.

1.1. Importance of foreign language reading

Publications in the field of foreign language teaching provide several arguments for the importance of reading in a foreign language. An early example of the recognition of the importance of reading is the work of West, discussed in Carpay (1975). In several publications between 1926 and 1960, West makes out a case for the teaching of reading. There are three types of arguments that underline the importance of teaching reading. Firstly, reading is an activity that plays an important role in our society (social-cultural argument). Secondly, it is easier to attain a high level of receptive language skills than of productive language skills (educational-psychological argument). Thirdly, relatively little preparation is required to organize a reading course and teachers can realize such lessons with relatively large groups of pupils (organizational argument) (cf. Van Esch, 1987).

Features of West's ideas are reflected in present arguments that underline the importance of foreign language reading skills. A difference can be made between reading as a teaching objective and as a tool to attain other objectives (cf. Van Esch, 1987).

Arguments in favour of reading as a *teaching objective* are mainly grounded on research of the practical needs for foreign language reading skills in society.

An early example of such research is the extensive investigation carried out in the Netherlands by Claessen et al. (1978). The outcomes showed inadequacies in

foreign language reading comprehension ability in professional, educational and everyday situations.

In a more recent publication Carrell (1988a) underlines the importance of teaching non-native reading skills, particularly in English as a second or foreign language. Carrell asserts that reading is especially important at an advanced level of non-native language proficiency and if needed for academic purposes.

Present developments in our society point out the needs for foreign language reading ability. Borders between countries are less of a hindrance to communication, particularly in the light of a united Europe. For many jobs, nowadays, it is a requirement to be able to read in a foreign language.

A relatively 'new' argument is the so-called 'digital revolution'. Nowadays many people have access to World Wide Web to search for information they are interested in. The amount of information available by means of this global multilingual network is growing at an incredible rate. This development contributes to the fact that the needs for foreign language reading skills are still growing.

One has to be careful to directly draw conclusions from these observed needs for teaching practice. The needs for foreign language reading skills, however, show the importance of teaching foreign language reading. They are an indication of what kind of reading skills students will have to develop and to what degree they will have to become proficient foreign language readers.

A second argument in favour of reading as a teaching objective, at least at the beginning stage of foreign language learning, is based on both theoretical and practical considerations. Theory about how people learn a foreign language reveals that it is easier to achieve a high level of reading and listening comprehension than of productive language skills. Considering the fact that the time assigned to foreign language teaching and learning is often limited, it is preferable to concentrate on the development of reading and listening and achieve a high level of receptive language skills, instead of dedicating this time to productive foreign language skills. Students will experience success and feel encouraged to continue their foreign language study (cf. Schouten-van Parreren, 1983; Staatsen et al., 1994).

Finally, there are some other practical arguments for reading as a teaching objective. Reading is pre-eminently suitable for teaching at differential levels of foreign language ability. Students can read individually, can choose their own reading speed and can read texts that fit their reading competence. Additionally, reading can easily be done outside the classroom as well and can be maintained and improved independently from established language courses (cf. Barnett, 1989).

There are also several arguments to stress the importance of reading as a *tool* to satisfy other foreign language needs.

Nowadays, it is generally believed that reading triggers the development of productive language skills, i.e. speaking and writing. Therefore, reading plays an important supporting role in communicative models of language teaching. In the ABCD-model (Neuner et al., 1981) understanding authentic texts in the first phase,

serves as a starting point for activities in the other three phases of the model (reproduction, production and transfer of productive language skills to other situations).

Reading as a tool has often been investigated with regard to specific aspects of foreign language acquisition, especially vocabulary. In several studies the importance of reading as a means of learning foreign language vocabulary in context is emphasized (e.g. Schouten-van Parreren, 1985; Mondria, 1996).

Traditionally, reading also plays a crucial role in the foreign language literature and culture curriculum. Often reading a foreign language is regarded to be essential to appreciating target language literature (cf. Barnett, 1989) and to gathering knowledge of and creating understanding in cultural and political aspects of the target country (cf. Staatsen et al., 1994).

1.2. Reading and foreign language teaching

In the previous section, several reasons were mentioned as to *why* it is important to teach foreign reading comprehension ability. The current section is about the role of reading in influential views on the teaching of foreign languages. The teaching of foreign language reading has been influenced by prevailing ideas about *what* is to be taught in the foreign language classroom and *how* this can best be done.

Views on foreign language teaching influence the objectives, classroom activities and choice of texts of foreign language reading. There are two approaches that, throughout time, have strongly influenced foreign language teaching, i.e. the Grammar-Translation Method and the Direct Method (cf. Van Els et al., 1984).

The *Grammar-Translation Method* is often associated with the teaching of foreign languages in the 19th century. At the time, the modern foreign languages were in competition with Latin and Greek. To strengthen their position they imitated the way the classics were taught. Reading in the traditional approach means translating, i.e. decoding, word by word, with the aid of a dictionary or word list and the help of the teacher, what a text is about. The Grammar-Translation Method focuses on teaching the target language's grammar rules and paradigms. Therefore, the texts that are used introduce specific grammatical problems. At a more advanced level, the texts are meant for developing students' appreciation of the target language's literature and culture.

The *Direct Method* allows students to perceive meaning directly through the target language because no translation is allowed. An important exponent, the *Audio-Lingual Method*, strongly influenced theory and practice of foreign language teaching in the sixties, as a reaction to the Grammar-Translation Method. The Audio-Lingual Method is based on the behaviourist belief that language learning is the acquisition of a set of correct language habits, not of a set of grammar rules. Teaching techniques concentrate on the production and understanding of spoken language. The role of reading, especially at the beginning of the Audio-Lingual period, is limited. Competing views on foreign language teaching - such as the

Reading Method developed by West in the United States in the 1930s, which came into use in Europe after World War II - never reached a comparable degree of popularity. In the Audio-Lingual Method, written materials serve as examples of spoken language (dialogues, conversations) or for practising language structures (pattern drills). Reading is considered to be a tool to learn to speak in the foreign language. Reading means reading aloud, concentrating on pronunciation, dedicating attention to language structures and unknown words. Texts serve for introducing or repeating grammatical structures and vocabulary, not for practising reading skills.

Currently, also a third approach towards teaching foreign languages has been adopted. This *Communicative Approach* was developed in the 1970s and is based on functional aspects, instead of theoretical assumptions underlying language teaching. According to the Communicative Approach the main objective of language teaching is 'communicative competence'. In other words, students have to learn how to use the foreign language in everyday situations. The central idea is that learning a foreign language above all means becoming a skilled language user. Therefore, the formulation of teaching objectives must be preceded by a thorough analysis of the situations in which a foreign language is to be used in practice. Firstly, it is important to establish what students have to achieve, in what situations they have to be able to communicate in the foreign language and how they must perform in these situations. Secondly, a further analysis of these situations is to provide insights into required linguistic *functions* (e.g. requests, denials, offers, apologies) and the general and specific *notions* or concepts a student will need to deal with in verbal communication (concepts such as time, sequence, quantity or concepts related to a specific topic, such as travelling or sports). Finally, an inventory is made of the linguistic structures and words that are required to perform these functions and notions (cf. CVE-MVT, 1992). In the beginning of the 'communicative era', most attention was paid to the development and teaching of productive language skills, especially spoken language. Later, the importance of reading as a communicative tool became increasingly well recognised (Neuner, 1980). Nowadays reading not only plays a supporting role, but it is considered to have its own specific communicative value. The main goal of reading is comprehension, not translating (Grammar-Translation Method) nor reproducing language structures (Audio-Lingual Method). Functional reading includes such activities as extensive reading, skimming and scanning, which require relatively little reading ability. It also includes the use of frequently consulted, authentic texts, such as travelling guides and instruction manuals. In a broader approach to communicative reading, the role of reading is not just limited to functional texts, but also to more complex text types (Van Krieken, 1984).

1.3. The Netherlands: Teaching practice

In the previous section, it was outlined that the role of reading in views on foreign language teaching has been limited. In the Grammar-Translation Method reading serves for getting familiar with grammar rules and literary or cultural aspects of the target language, whereas in the Audio-Lingual Method reading is a tool to learn to speak in the target language. The current teaching practice of foreign language reading in the Netherlands still embraces aspects of both views. Often teachers focus on knowledge aspects of reading: they explain unknown words and check what pupils have understood of a text. There is little attention for the developing of skills that pupils actually help to understand texts (Westhoff, 1989).

Additionally, the way reading is tested has an impact on teaching practice in the Netherlands. This is the so-called *backwash-effect*. In the Netherlands standardized reading comprehension tests are used for the final examinations of the foreign language courses in secondary education. These tests are developed by CITO (National Institute for Educational Measurement) and are compulsory for all pupils in secondary education. They consist of a series of reading texts with items in multiple-choice format. As a consequence, in the foreign language lessons, pupils often labouriously practise (multiple-choice) questions about reading texts, in preparation for the final examinations. According to Westhoff (1989), this hardly leads to better reading comprehension because teachers only check if pupils have understood text contents, whilst placing little emphasis on the developing of reading skills.

The Communicative Approach has changed the perspective on the objectives of foreign language reading. The outcomes of research into the needs for foreign language skills show that in many situations reading skills are required. Reading is considered to be a language skill with its own specific objectives. These objectives, however, are directly based on established needs for language skills, not on insights into how students learn to read in a foreign language. For teaching practice also cognitive aspects of language learning are important. Characteristic for this cognitive-psychological view is that learners process the same input in different ways, leading to differences in the output. Important variables in this process are variation in human processing, the nature of the (reading) task and differences between individual learners. This means that, besides linguistic aspects of reading (vocabulary, grammar), attention should be paid to differences between students and to reading techniques and strategies. As Barnett (1989) says in the introduction of her study on foreign language reading "... we see reading in a different light: as communication, as a mental process, as the reader's active participation in the creation of meaning, as a manipulation of strategies, as a *receptive* rather than a *passive* skill." (p.2).

In the next chapter cognitive aspects of reading will be comprehensively discussed from a theoretical point of view. In the Netherlands, in the recent reforms in the

curriculum of secondary education both communicative needs and cognitive learning theory have played a part. This affects the teaching of foreign language reading. Reading objectives reflect the needs for foreign language reading skills (communicative view), as well as insights into the mental processing involved in learning and reading (cognitive view). The next section is about these reforms. The emphasis is on the teaching and learning of foreign language reading comprehension.

1.4. The Netherlands: Recent reforms

Except for some elementary English training in primary education, foreign language teaching takes place during secondary education. Since 1968, this secondary education has been officially subdivided into a first and second phase. Secondary education in the Netherlands is for pupils aged 12 and over. It is divided into pre-vocational education (VBO), junior general secondary education (MAVO), senior general secondary education (HAVO) and pre-university education (VWO). Pre-vocational education (VBO) and junior general secondary education (MAVO) both last four years, senior general education (HAVO) lasts five years and pre-university education (VWO) six years. The recent reforms in secondary education are the introduction of basic secondary education in the first phase of secondary education, changes in the top years of VBO/MAVO and the innovations in the second phase of HAVO and VWO.

During the first phase of secondary education pupils are taught a compulsory core curriculum, referred to as *basic secondary education*. This basic secondary education was introduced in 1993 and includes at least two foreign languages: English, which is obligatory for all pupils, and French or German. Depending on the type of education (VBO, MAVO, HAVO or VWO), this core curriculum is extended by a deepening of the subjects and by other subjects of instruction.

Following the period of basic secondary education, pupils spend the second phase of secondary education preparing for the school-leaving examinations. These examinations are produced and administered under central government supervision. From 1 August 1999, both VBO and MAVO will be replaced by the new VMBO (pre-vocational secondary education). This new VMBO will be composed of four sectors: engineering and technology, economics, agriculture and care and welfare. Within each sector there will be a choice of three learning pathways: theoretical, vocational (available at two levels) and combined theoretical and vocational. For each pathway there will be a fixed combination of examination subjects.

From 1 August 1998, there will be changes in the second phase of the HAVO/VWO level. Four set subject combinations will be introduced, replacing the present more or less free choice of subjects. These set combinations are science and technology, science and health care, economics and society, culture and society. Each of the four set subject combinations consists of a common, a specialised and an optional component. The common component provides a broad general

education, will absorb almost half of the available classroom time and is the same for all set combinations.

The main developments in secondary education relate to teaching itself. The new concept of schools as 'places of study' means that there will be a greater emphasis on autonomous learning. Three interconnected aims can be distinguished: to encourage the broad personal development of all pupils, to further an active and autonomous attitude towards learning and to provide a teaching environment in which differences between pupils are recognized. The reforms are intended to smooth the transition from secondary to higher education and are being coordinated by a secondary education process management team, created for the purpose by the government (OC&W, 1998).

The teaching of foreign languages in the innovated curriculum is based on a communicative approach, cognitive learning theory and theories about autonomous learning. The emphasis is on the development of foreign language skills; knowledge of the linguistic system is considered to be less important. In case of deficits in language proficiency, pupils have to master compensatory strategies to satisfy their communicative needs. Compensatory strategies embrace all kinds of strategies and techniques that can help to increase a pupil's communicative competence.

This different approach towards language teaching affects the teaching of foreign language reading. Secondary education aims at a higher level of reading comprehension ability than of listening comprehension, speaking and writing. This is based on the cognitive insight that it is easier to attain a high level of receptive than of productive language skills and the estimation that it is advisable and useful, for further study and future jobs, to teach foreign language reading comprehension ability. Teaching foreign language reading implies focusing on the process of reading comprehension, i.e. the steps that can be taken to achieve a certain reading goal. In other words, pupils have to learn to apply reading strategies (techniques, tools and methods of approach of a text or a text part). Teaching reading skills also embraces the use of compensatory strategies, such as contextually guessing the meaning of unknown words and making use of a dictionary or a survey of the foreign language grammar. This way of teaching foreign language reading in secondary education is very different from how reading has usually been dealt with (see 1.3). Besides asking questions to check whether pupils have understood a text, the teaching of reading aims at helping pupils to systematically improve their reading comprehension skills. In other words, they actually *learn* to read in a foreign language. They acquire reading skills by 'doing things with texts', e.g. establishing what type of text they are dealing with (a leaflet, a newspaper article etc.), predicting what a text is about and searching for words that are familiar to them. Performing this type of tasks teaches pupils to cope with future reading activities.

The use of reading and compensatory strategies is a constant theme running through the objectives of foreign language reading in secondary education, such as formulated in the innovated curriculum. The activities that can be organized to teach

reading strategies in the foreign language classroom belong to three sections. Activities of the first section refer to predicting and verifying (e.g. predicting contents by means of subtitles, recognizing a text type by means of its unique features). Activities of the second section are exercises that make pupils aware of the structure of a text (e.g. establishing the right order of a set of loose text parts, inventing a subtitle for each paragraph of a text, making a diagram of text contents). The third section embraces guessing and making use of redundant information (e.g. contextually guessing, gap-filling exercises). For each of these sections, examples of activities to develop reading skills have been provided.

There is also differentiation between the reading objectives of the different types of education. This is reflected in the degree of comprehension of a text required and the text types to be used as teaching material. Below, the specific reading objectives for basic secondary education are discussed and the reading objectives for the second phase of VBO/MAVO (the future VMBO) and HAVO/VWO.

To what *degree* a pupil has to be able to understand a text depends on the phase of secondary education (first or second phase) and the type of secondary education.

The core curriculum of basic secondary education of the first phase embraces five objectives for foreign language reading. According to the first objective, pupils have to show that they understand directions, superscriptions, warnings and announcements. The second and third objective refer to the reading of so-called functional texts. Pupils must be able to look for relevant information and to match this information in order to draw conclusions. Corresponding to the fourth objective, pupils have to demonstrate that they understand the main points of informative texts. Finally, pupils have to gain experience with respect to the extensive reading of longer texts (Staatsen et al., 1994).

In the second phase there is a distinction between two broad proficiency levels: the VBO/MAVO and HAVO/VWO-level. At both levels, pupils have to be experienced *extensive* readers. This implies that they have to read an ample number of easily comprehensible texts. It is expected that, by means of this receptive input, pupils will acquire knowledge of foreign language vocabulary and syntax. The exigencies for *intensive* reading are different for each of these two levels. Pupils of VBO/MAVO have to demonstrate that they are able to understand essential information - the main points and important details - of (practically) authentic texts. Pupils of HAVO/VWO have to meet more specific criteria. Firstly, they have to be able to judge whether or not a certain text contains information that is relevant to them. Secondly, they have to understand the main lines of text contents. Thirdly, they must understand some (relevant) text parts in detail. Fourthly, they have to be competent to distinguish between text parts and to establish the connections between them. Finally, they have to be able to draw conclusions from a text with respect to, for example, the author's target group, his intentions and views. Besides these five reading objectives, there are two general objectives related to foreign language reading. Firstly, pupils have to know how to make a summary of a text and use this skill as a strategy to deal with a text. Secondly, pupils have to know how to use

modern information technology to gather information (e.g. CD-rom and World Wide Web).

A second point that is elaborated in detail is the *choice of texts*.

The texts that are suitable for use in the programme of basic secondary education belong to four broadly defined text types: superscriptions and warnings, functional texts (e.g. brochures, advertisements), informative texts (e.g. newspaper-articles) and larger texts, including fiction. The criteria for the choice of texts are the following. Text structure must be rather simple, the texts must be written in comprehensible words and the main information must be explicitly mentioned. The topics have to agree with and build on the pupils' own experiences and levels of maturity. Finally, the materials must resemble materials from everyday life. This means that texts either have to be authentic or must have the characteristics of authentic texts (contents and lay-out) (Staatsen et al., 1994).

In the second phase texts used at the proficiency level VBO/MAVO have to meet the two following criteria: they have to cover topics that link up with the pupils' own interests, their futures, and social events they are interested in and the main points and main details of the texts must be understandable, taking into account the pupils' knowledge of the foreign language vocabulary and structures. Examples of texts that meet these criteria are informative letters, leaflets, written instructions, texts from newspapers, magazines, or books, public announcements, advertisements and the menu of a restaurant (O&W, 1986; Staatsen et al., 1994).

There are somewhat different criteria for the choice of texts at the HAVO and VWO-level. For each of these types of education there are criteria describing in detail the level of familiarity of the text topic, the prior knowledge of target culture needed to understand a text, the complexity of text structure and the way the information in the text is presented (cf. VOG-MVT, 1995).

The Dutch reading objectives reflect a communicative view on foreign language teaching which takes into consideration the needs for foreign language skills in society. In this view, both extensive and intensive reading are considered to be of importance. A strong point is that these objectives emphasize the development of reading skills and the importance of autonomous learning. It is recognized that pupils have to be familiar with strategies and techniques that help them to understand texts and that there are differences between pupils in how they approach a reading task.

Although the foreign language reading objectives such as formulated in the innovated curriculum for secondary education in the Netherlands are far more specific and detailed than in the previous plan, there are also points of criticism.

Struik (1996) believes that the criteria for the different proficiency levels of secondary education are not easy to apply in practice.

Pit (1996) discusses the criteria for the choice of reading texts. In the new programmes, these criteria focus on aspects such as text topic, text structure and the way the information is presented (e.g. level of redundancy, explicitness), instead of

linguistic criteria in a restricted sense (vocabulary knowledge, grammatical structures). According to Pit this is not an improvement. Just as in the case of linguistic criteria, the choice of texts according to the new curriculum will often have to be based on intuition. Pit believes that criteria for text choice have to embrace linguistic aspects as well. Pit reports the outcomes of recent findings of a study carried out by CITO (cf. Pit & Van Zuylen, 1996), which show that pure linguistic criteria are among the best predictors of the level of difficulty of reading texts.

1.5. Summary and conclusion

This first chapter dealt with the status quo of methods and objectives of teaching foreign language reading. One of the conclusions that can be drawn is that reading is an important foreign language skill. Reading fulfills needs in society, is a means of language acquisition and motivates students to learn a foreign language.

Firstly, the role of reading in foreign language teaching was discussed. Traditional views on foreign language teaching mainly focus on reading as a tool to achieve other teaching objectives. In the Grammar-Translation Method, reading is a means to learn the foreign language grammar and in the Audio-Lingual Method reading serves for learning to speak in the foreign language.

Secondly, some characteristics of foreign language teaching in the Netherlands were dealt with. Teaching practice still shows aspects of traditional views on foreign language teaching and there is a backwash-effect of the central school-examinations produced by CITO.

Finally, current developments in the field of teaching foreign language reading were outlined. The Communicative Approach has changed traditional views on the teaching of foreign language reading, stressing the importance of such aspects as the functional character of reading and working with (semi-) authentic texts. For the practice of foreign language teaching, however, also cognitive aspects of language learning are important, i.e. insights into how people learn and insights into the process of foreign language reading. The recently updated foreign language reading objectives for secondary education in the Netherlands are an example of teaching objectives that reflect these insights. The Dutch objectives emphasize autonomous learning and embrace the teaching of reading and compensatory strategies. This is particularly important in the view of this study, the central question of which is whether context-use training is useful to develop reading skills. The answer to this question is even more relevant now that strategy use has become an official objective of the teaching of reading in secondary education in the Netherlands.

Before trying to answer the central question of this study, it is necessary to know more about the mental processes underlying reading. This will be the topic of the next chapter.

Chapter 2

Reading

Developing models about how readers understand written language is a relatively 'new' research area. Although we know of earlier examples of reading research, e.g. the paper about letter and word recognition by Cattell (1886), it was not until the mid-1960s that explicit models of reading were developed, that is, "models that describe the entire process from the time the eye meets the page until the reader experiences the 'click of comprehension'" (Samuels & Kamil, 1984/1988, p. 22). This burst of model-building activity was a direct consequence of the impact of cognitive science on reading research.

From the 30-s till the 60-s, in the Skinnerian era, 'behaviourism' strongly determined psychological theory and research, also in the field of language learning. The study of human mental processes was considered to be both methodologically impossible and irrelevant. This changed when cognitive science was born in the 60-s. The question of how learners approach the different tasks of language learning gained interest and the study of the mental processes of learners came to occupy a central role in language research.

Nowadays, this cognitive perspective of language learning has an enormous impact on reading research, resulting in different views and models. In addition, studies in the field of artificial intelligence have contributed valuable insights.

In this chapter current views on information processing and reading comprehension are discussed. In section 2.1 three models of reading are outlined: top-down, bottom-up and interactive processing. Nowadays it is generally accepted that reading is an interactive process in which both bottom-up and top-down processing take place. Therefore, in section 2.2, three aspects of the interactive view are discussed, that is, knowledge-based, text-based and compensatory processing. Section 2.3 is the conclusion of this second chapter.

2.1. Three models of reading

It is difficult to find a model of reading by means of which all processes involved can be explained. Firstly, the course of a reading activity always depends on various factors. Each individual reader will process a text differently, given his cognitive development, personal reading skills, previous knowledge of text topic and the purpose of reading. Secondly, a single, generally accepted theory of reading is not available. Researchers differ in their theories and models of reading with regard to the relative importance they contribute to the processes involved. Models reflect

personal views on reading of individual researchers and the trends and issues in reading research of the period in which they were developed. Therefore, models are rather partial than complete and difficult to compare, because they often explain different aspects of reading (Samuels & Kamil, 1988; Barnett, 1989).

Often researchers make a distinction between two broad processing levels: processes at the *lower* cognitive levels and processes at the *higher* cognitive levels (e.g. Perfetti, 1985; Stanovich, 1991). Lower-level processing means decoding the graphic signs (letters, words) of a text, whereas higher-level processing signifies the complex task of making sense of this graphic input. The terms 'higher' and 'lower' are used in a figurative sense and are based on the assumption that reading comprehension takes place at different cognitive processing levels that can be ordered in an hierarchical way: from low-level recognizing graphic input to high-level text comprehension or vice versa.

Reviewing studies on reading comprehension, Segalowitz et al. (1991) specify the processes belonging to each of these two levels. Lower level-processing includes the visual recognition and identification of letters, the generation of grapheme-phoneme correspondences, the use of orthographic redundancies and the association of words with their semantic representations. Concerning higher level processing they refer to aspects such as resolving ambiguities in the text, integrating textual information with prior knowledge and updating what is understood of the text as a whole.

At first, a very rigid interpretation of the hierarchy of processing levels generated two groups of reading models, bottom-up and top-down. These two views on reading are outlined below.

Bottom-up models explain reading as a decoding process, starting at the 'bottom', i.e., the identification of the graphemes of the text. Bottom-up models are strictly text-based and do not allow any higher-level process to affect the processing at a lower level. A reader processes serially, letter by letter and word by word, just as much print as primary, short-term memory can handle. When all words have been identified, the next chunk of text is processed in exactly the same way. Bottom-up models do not explain higher-level processes. No clear explanation is provided of how readers, after having recognized and given meaning to the printed letters and words, eventually understand sentences and larger text parts (cf. Barnett, 1989). To solve this problem Gough (1972) even had to invent a kind of miraculous device, named 'Merlin', by means of which sentences in primary memory are interpreted and then stored in what he called the PWSGWTAU (= Place Where Sentences go When They Are Understood). Other examples of bottom-up models are the models described in LaBerge & Samuels (1974), Carver (1977-78) and Just & Carpenter (1980).

Top-down models explain reading as a process that starts at the 'top', at the higher-level interpretative processes. Top-down models are strictly knowledge-based: a

reader uses prior knowledge to predict or guess the contents of the text and samples just enough text to confirm or reject these hypotheses. The most-cited example is the model of Goodman (1967; 1973), although Goodman himself denies that his model is a top-down model (Goodman, 1981). Goodman defines the process of reading as 'a psycholinguistic guessing game' consisting of predicting, sampling, confirming and correcting. Smith (1971) also reflects on reading as a top-down process and, without really building a reading model, adds the notion of purposeful reading. According to Smith, readers sample text not only to confirm or reject predictions of just any kind, but also to reach a specific reading goal. Often people read texts for a certain reason and only pay attention to clues in the text that are related to this purpose.

Both bottom-up and top-down models are too rigid to consider all aspects involved in the reading process. Bottom-up models do not account for the facilitating effects of context and prior knowledge on the perception and identification of the letters and words of a text and on the interpretation of words and sentences (Rumelhart, 1977; Stanovich 1980). Strictly top-down models, on the other hand, hardly take into account the information provided by the text itself. Describing reading as predicting text contents rather than processing the graphic input, strictly top-down models strongly overestimate the importance of higher-level processes (Stanovich, 1980; 1991).

Therefore, strictly bottom-up and strictly top-down models were soon replaced by interactive models, which no longer define reading as a serial bottom-up or unidirectional top-down process, but explain how lower-level and higher-level processes work together interactively. Interactive models include the strong points of both bottom-up and top-down models. They allow higher-level processing to influence lower-level processing and recognize the importance of information provided by print. Higher-level predicting and interpreting reduce the possibilities at the lower sentence or word levels, while at the same time these lower-level decoding processes limit the possible number of interpretations (cf. Aarnoutse, 1982). Nowadays, this interactive view is generally accepted by reading researchers.

2.2. Aspects of interactive reading

In the interactive approach, reading is characterized as the process of 'creating' comprehension while interacting with print (Rosenblatt, 1978; Widdowson, 1979). This means that a text itself has no meaning, but only 'a potential for meaning' (Widdowson, 1979). Eskey (1988) defines interactive reading as follows "... *interactive* refers to the interaction obtained by means of bottom-up decoding and information provided by means of top-down analysis, *both* of which depend on certain kinds of prior knowledge and certain kinds of information processing skills" (p.96). Eskey's definition shows that interactive views on reading postulate that several processes at different levels are happening at once and each process can affect processes at other levels.

In interactive models of reading much attention is paid to processes of inference. Strictly speaking, the term 'inference' refers to the logical conclusions one can draw about information that is not explicitly mentioned. For example, from the information *Marc is taller than Marcel* and *Marcel is taller than Peter*, the following logical inferences can be made: *Marc is taller than Peter* and *Peter is the shortest*. Often, however, the term 'inference' is used in a broader sense and also embraces interpretations that are probable, possible or plausible (cf. Aarnoutse, 1982).

Below, examples of models of interactive reading are discussed. Some of these models emphasize the role of prior knowledge, others the role of graphic input. A third category is the interactive-compensatory view, which claims that processing at any level (lower or higher) can compensate for deficiencies at any other level.

2.2.1. Knowledge-based processing

In many models of interactive reading the role of prior knowledge is stressed. Higher-level predicting and interpreting direct reading, reduce the possibilities at the lower sentence or word levels and facilitate word recognition, while at the same time these lower-level decoding processes support higher-level comprehension and limit the possible number of interpretations. In many studies, these models are referred to as *schema-theoretic models*.

The term 'schema' was first used by psychologist Bartlett in his classic study *Remembering* (1932) and has been further developed by other researchers in the field of cognitive science, including reading research. A schema is an abstract knowledge structure stored in human memory. Schemata represent the reader's prior background information about concepts, situations, facts, topics, actions or events. Schemata can be applied in every imaginable context, e.g. watching T.V., planning a holiday, ordering soup in a restaurant, identifying a dinosaur egg, travelling by aeroplane, building a house, investigating a murder, reading a magazine and writing a thesis. A reader can use this information to create a context for the graphic elements of a text. Rumelhart & Orthony (1977) give four basic characteristics of schemata.

The first characteristic is that schemata represent generic, abstract knowledge. Schemata summarize what is generally considered to be true about objects, situations, events, actions and the sequence of events and actions. There is a difference between a person's knowledge tied to specific instances (e.g. last night's dinner with John in the Italian restaurant *Roberto*) and his schemata, i.e. the abstractions of situation-specific experiences to general concepts associated with a certain type of situation (e.g. eating in a restaurant). Schemata symbolize knowledge at different levels of abstraction, from basic perceptual elements such as the configuration of lines which form a square, to complex networks of related concepts such as a 'restaurant schema' involving actions (e.g. entering, ordering, eating, paying, leaving) and concepts (e.g. menu, waiter, food, bill) related to the situation 'eating in a restaurant' (cf. Schank & Abelson, 1977).

Another characteristic of schemata is that they represent conceptual knowledge, i.e., knowledge that is not stored in memory in the form of words. Consequently, schemata are not related to a specific language

The third characteristic is that schemata can include each other. Rumelhart & Orthony (1977) give the example of the schema of the 'human face', in which is embedded the sub-schema of the 'eye'. The latter, in its turn, includes the sub-schemata of 'pupil', 'iris' and 'eyelid'

The fourth characteristic is that schemata contain variables - also named 'nodes' or 'slots' -, i.e. default values which represent a large number of individual cases. In the example of the 'restaurant-schema' one of the slots is the restaurant itself. This slot can be filled in by any specific restaurant a person knows: the restaurant at the corner, the nice little restaurant in Greece or Italian restaurant *Roberto* where we dined with John

Much work in reading research concentrates on the relevance of schemata in the process of reading comprehension (for a review, see Anderson & Pearson, 1984/1988). Often a distinction is made between prior knowledge of the content area of a text (content schemata) and formal, rhetorical organizational structures of different types of texts (formal schemata) (Carrell, 1983). It is postulated that, during reading, schemata are activated and that this information contributes to the interpretation of text contents. The new information of the text either fits in one of the reader's schemata or the reader modifies existing schemata in such a way that this new information is integrated. Or, in the words of Anderson & Pearson (1984/1988): "To say that one has comprehended a text is to say that she has found a mental 'home' for the information in the text or else that she has modified an existing home in order to accommodate that new information" (p.37). Schema theory has strongly influenced views on interactive reading, because it explains the facilitating effects of higher-level interpretative processing on reading comprehension. But there are also points of criticism. Schema theory sticks at the theoretical level and cannot explain how the actual activation and updating of schemata works, nor what is the relative contribution of schemata to reading comprehension. Two examples of schema-theoretic models of reading are discussed below, i.e. the Rumelhart model (1977) and the Anderson & Pearson model (1984/1988).

The Rumelhart model

Rumelhart (1977) is one of the first to recognize that reading is not a linear bottom-up process, but interactive. He theorizes that reading comprehension means selecting schemata and plots which fit best the input of the text. According to Rumelhart, a reader uses all available prior knowledge sources, i.e. semantic, syntactic, orthographic and lexical knowledge, simultaneously to extract information necessary for understanding print. The graphic input is first registered in a visual information store (VIS) and then checked by the feature extraction device, pulling out the critical features of the input. The outcome of these processes go into the pattern synthesizer, where all the reader's previous knowledge about

language spelling patterns, syntax, vocabulary, semantics and context work together to interpret the graphic input. A higher mechanism, the message centre, coordinates the processes that take place in the pattern synthesizer. It temporarily stores hypotheses about the input from each of the four knowledge sources, permits each knowledge source to search for relevant hypotheses from other knowledge sources, it removes rejected hypotheses and adds new ones, until an agreement is reached about the most probable interpretation.

A strong point of Rumelhart's model is the fact that it explains that higher-level information interacts with the analyses at lower levels and, thus, accounts for a number of observations in reading research, which, at that moment, could not be properly explained.

Firstly, Rumelhart's model clarifies that the processing of non-words or nonsense words takes more time than the processing of existing words containing the same number of letters (e.g. the word 'alligator' versus 'rllaagtio').

Secondly, the model explains that semantic information influences word perception. For instance, semantically related word-pairs, such as 'bread-butter' or 'doctor-nurse' are processed faster than not related word-pairs, like 'bread-nurse' or 'doctor-butter'.

In addition, it explains that perception of syntax is highly related to context. For example, the sentence "They are eating apples" is ambiguous, because it is not clear whether it refers to the act of consuming apples or to a type of apple: an 'eating apple' as opposed to a 'cooking apple'. If such a sentence is preceded by the question "What are the children eating?" or "What kind of apples are these?", it will probably not lead to misunderstanding. Finally, the model developed by Rumelhart accounts for the facilitating effects of context on the interpretation of larger text parts. In both of the following sentences the word 'figures' appears: (1) "The statistician was certain the difference was significant since all the figures on the right-hand side of the table were larger than any of those on the left"; (2) "The craftsman charged more for the carvings on the right side since all the figures on the right-hand side of the table were larger than any of those on the left". The interpretation of the meaning of 'figures' in the first sentence is that of numerals and in the second sentence of a wooden or ceramic figure. This example shows that we do not just extract meaning from the text segment we are processing but also from its surrounding information (cf Samuels & Kamil, 1984/1988).

The original Rumelhart model (1977) was later expanded by Rumelhart and some colleagues (Rumelhart et al., 1986). With the aid of computer simulations, they constructed so-called PDP (Parallel Distributed Processing) models, which explain how the mind processes information. The model not just refers to processing of written information, but to information processing in general. As in the original Rumelhart model (1977), information processing is described as the interaction of many different processing elements which all work at once. In the updated model (Rumelhart et al., 1986) these elements are referred to as units. Units may contain hypotheses (e.g. about letter sequences or the syntactic function of words in a sentence) and send excitatory signals to units containing connected hypotheses and inhibiting signals to units with hypotheses that do not fit. The mind considers all

these options in order to comprehend a text.

The Anderson & Pearson model

The study of Anderson & Pearson (1984/1988) is not really a model but an exploration of one single aspect of interactive reading, that is, the working of content schemata. They based their study on earlier research in the field of schema theory and reading comprehension and added some new insights.

Anderson & Pearson (1984/1988) review what they call "a simple model of schema-activation" (p.44), representing what is generally assumed about the working of schemata. They use as an example the schema of the christening of a ship, including a number of slots (e.g. 'done by a celebrity', 'bottle broken on bow'). They explain that it is assumed that, when the schema gets activated, the slots are 'filled in' with particular information. Such information is constrained. It is likely that the 'celebrity' slot is filled in by a congressman or the Prince of Wales, but not by a garbage collector or barmaid. It is also assumed that some components of a schema are more likely to activate the whole schema than other components. In the case of the ship-christening schema, the slot 'bottle broken on a bow' is more likely to remind a reader of a ship christening than the 'celebrity' slot, because the latter could also be part of other schemata representing knowledge of situations in which a celebrity takes part. A final assumption is that the more components of a particular schema are mentioned, the more probable it is that this schema is indeed activated. They illustrate this by the following two sentences: (1) "Princess Anne broke the bottle on the ship"; (2) "The waitress broke the bottle on the ship". In the first sentence three slots of a ship-christening schema are mentioned ('celebrity', 'bottle-breaking' and 'ship'), whereas in the second sentence only two ('bottle-breaking' and 'ship'). Therefore, the first sentence is more likely to be interpreted as a ship christening than the second one.

Anderson & Pearson (1984/1988) argue that this simple, generally accepted, model of schema-activation needs to be developed further. The model is not able to explain why the sentence "During the ceremony on the ship, Prince Charles took a swig from the bottle of champagne", which also contains a number of slots that may refer to a ship christening, is presumably not interpreted as such. According to Anderson & Pearson, this can be explained by the fact that readers make inferences about logical relationships between the components of a schema. These inferences determine which information matches a schema and which information does not. The Prince Charles sentence will probably not be connected to a ship-christening schema because of inferences such as: "If the celebrity's taking a drink from the bottle were a part of a ship christening, I would probably know that fact" and "A ceremony during which the celebrity takes a drink from a bottle is probably not a ship christening". They conclude that an adequate description of the structure of schemata will include information about logical relationships among components and that a complete theory of schema-activation will include a major role for inferences.

Besides schema-activation, Anderson & Pearson (1984/1988) discuss a second aspect of schema theory, i.e. the interaction between abstract knowledge embodied

in schemata and memory for particular examples. Anderson & Pearson assume that the role of the latter is more important in reading comprehension than conventional schema theory postulates. They state that it is not unthinkable that "much that passes for general knowledge is actually derived as needed by retrieving specific cases and making calculations based on what is known about them" (p 51). They discuss a well-known theory in the field of schema theory elaborated by Collins & Quillian (1969) who propose the idea of 'cognitive economy', assuming a hierarchical structuring of concepts (e.g. a canary is a bird, a bird is an animal etc.) in which all characteristics that are valid for the concept 'animal' are valid for the concepts 'bird' and 'canary' as well. The greater the distance in the tree that has to be traversed to find the information, the longer it will take to interpret incoming information. Anderson & Pearson, however, argue that this hierarchical structuring is reasonable only for situations that do not belong to personal experience. For example, most people have probably never seen a giant condor fly, so there is no reason to assume that this information would be directly stored in their condor-schema. But in the case of a robin this situation is probably different. The average person has seen many flying robins. Therefore, it is more likely that the information that a robin can fly is directly stored in a person's robin representation, instead of at a higher level of the bird-schema.

Additionally, Anderson & Pearson (1984/1988) refer to a number of examples from reading research in which knowledge of particular cases appears to overrule the generic abstracted knowledge stored in schemata. For instance, Half et al. (1976) show that the red visualized in the following compounds 'red strawberry', 'red barn', 'red sunset' and 'red hair' is different in each of these cases. Anderson & Pearson propose that a person retrieves specific memories of, for instance, red hair and uses these memories to establish the type of red that combines with this concept. They believe that the interpretation of the compounds is predicted on the basis of the generic concept of red and the generic concept of the object (barn, sunset and so on) only when the person has not experienced this combination before or when an indeterminate range of shades of red is possible.

Considering what is outlined above, Anderson & Pearson (1984/1988) conclude that, during language comprehension, people probably rely on knowledge of particular cases as well as on abstract and general schemata. The term 'particular case' not necessarily refers to the memory of a case experienced at a particular moment, but can be an abstraction as well. They explain this as follows: "A robin is a specific example of a bird, but notice that ROBIN is itself an abstracted and generic schema. Still more specific is the-robin-I-saw-nesting-in-the-hawthorne-tree-outside-my-front-door-this-morning" (p 52). Following Smith & Medin (1981), Anderson & Pearson (1984/1988) assume that people have knowledge represented at various levels of specificity.

2.2.2. Text-based processing

Other interactive models are text-based and regard reading comprehension as a process that is primarily driven by the graphic input of the text. Lower-level decoding directs reading and, at the very beginning of the process, reduces the

possible number of interpretations from higher-level interactive processing

A well-known model of this category is the model developed by Kintsch & Van Dijk (1978). This model is based on a propositional method. A proposition is a symbolic representation of the relationship between two or more concepts that either can be correct or wrong. By means of propositions, connections can be made within a schema or between schemata. An example of a proposition is BELOW (CHAIR, BALL), which could possibly mean 'the ball is left below the chair'. The propositional model of Kintsch & Van Dijk (1978) concentrates on the semantic structure of a text and explains reading comprehension as the result of two reduction processes: the forming of text-related micropropositions and the forming of abstract macropropositions including only the most important information of a text. More recent interactive text-based models emphasize the role of word recognition in the reading process and are often based on the empirical results of eye-movement experiments (Just & Carpenter, 1987, Rayner & Pollatsek, 1989). It is postulated that efficient word recognition is a necessary condition for good comprehension. Following Fodor (1983), Stanovich (1991) characterizes word recognition of skilled readers as an autonomous (or *modular*) process. The Kintsch & Van Dijk and Stanovich model are reviewed below.

The Kintsch & Van Dijk model

Kintsch & Van Dijk (1978) describe text processing as the forming of text-related micropropositions and abstract macropropositions.

Micropropositions are semantic units that more or less represent a sentence or phrase of the text. During the forming of micropropositions, 'arguments' and 'predicates' play a major role. An argument represents the meaning of an individual word, in such a way that synonyms are represented by the same argument. A predicate (or 'focus') is a special argument. It connects the different arguments within a proposition. It represents the main argument and is mentioned first in a visual representation of a proposition. For example, (TEACH, SPEAKER, STUDENT) means that the predicate is teach. A predicate is not necessarily a verb, but can be any kind of word class. In the next example, it is an adjective: (VIOLENT, ENCOUNTER).

Kintsch & Van Dijk (1978) try to provide an explanation of the coherence of a text. They assume that text coherence is achieved by overlap of arguments in propositions. Therefore, they believe that a reader will constantly try to connect propositions which contain the same argument. These connections form the microstructure. The result of the process of forming micropropositions can be reflected in a 'coherence graph', a graphic representation of the microstructure.

Micropropositions are not exactly the same as the microstructure. The latter is the first abstraction of the list of micropropositions made with reference to a text and can be considered to be a relatively superficial representation of text contents.

The macrostructure, in its turn, is the strongly reduced version of the microstructure and is stored schematically in long term memory by means of macropropositions.

In the model of Kintsch & Van Dijk (1978), text processing is described as the

repetition of the same 'cyclic process'. a certain period of time in which a certain number of micropropositions are connected to each other and short-term memory is emptied, except for a handful of key propositions remaining there to be connected to the following gathering of propositions. This implies that, building the microstructure of a text, some propositions are considered to be more important than others.

Parallel with the forming of the microstructure, the macrostructure of a text is built, connecting important information of the microstructure with already existing knowledge structures, or schemata, in long term memory. A schema is filled in with selected propositions drawn from the microstructure. One goes from the microstructure to the macrostructure by deleting less relevant propositions, generalizing the information of different propositions and adding inferences.

As most reading models, the Kintsch & Van Dijk model does not explain the entire process of reading comprehension. It is not clear how words are recognized, how micropropositions are formed and how processes of inference come about (cf Rayner & Pollatsek, 1989)

In a second version of the model, published in 1983, Kintsch & Van Dijk try to improve the incompleteness and vagueness of their model. Rayner & Pollatsek (1989), however, do not consider this a real improvement, "because it substitutes vagueness for oversimplification" (p.290). The second version of the Kintsch & Van Dijk model has exerted far less influence than the earlier one.

The Stanovich model

Stanovich's study (1991) is a review of the role of word recognition in reading. Most recent models of the reading process stress the importance for reading comprehension of efficient word recognition, postulating that successful word recognition forms the foundation of, or is even a prerequisite for, higher-level comprehension processes (see e.g. Just & Carpenter, 1987; Rayner & Pollatsek, 1989).

According to Stanovich, word recognition, or *lexical access*, implies "the activation of the information associated with the word's orthographic representation, including the word's meanings and its phonological representation" (p.442). Based on a theory developed by Fodor (1983), Stanovich assumes that word recognition of skilled language readers is modular. A modular process operates without the interference of other processes and is, consequently, faster and more accurate than interactive processes. Fodor refers to this as "information encapsulation", that is, the situation where prior knowledge does not influence the process in question. As a consequence of modularized word recognition, readers do not need to 'waste' valuable processing time to identify and decode letters and words, but can save this for processes of interpreting and understanding at levels that go beyond the word level.

Stanovich explains that the value of this model is that it provides a simple explanation of the apparent contradiction revealed in empirical research that skilled readers, who had demonstrated the use of context skills, did not use context to

recognize words. In Stanovich's view, these readers prefer the efficient, modular process of word recognition to compensatory time-consuming processes such as the use of contextual information. Additionally, the model explains why poor readers are often more sensitive to contextual information than strong readers. In the case that word recognition skills are less automatized, a reader necessarily depends more on contextual information.

According to Stanovich, word recognition is an independent predictor of reading comprehension. He underlines that this does not imply that word recognition alone is sufficient for good comprehension. Comprehension also requires competent interactive processing at higher levels. Stanovich states that it is quite possible that a reader has poor comprehension ability despite adequate word recognition skills, but that it is very unlikely that excellent reading comprehension will be observed in a reader with weak word recognition skills.

2.2.3. Compensatory processing

The Stanovich model, such as outlined above, represents only part of Stanovich's ideas about interactive reading. Stanovich's complete theory not only accounts for the efficient comprehension process of highly skilled readers, who possess strong or even encapsulated word recognition skills, but also explains less skilled reading. Stanovich (1991) hypothesizes that the information processing system works in such a way that when a reader's word recognition skills are deficient, he or she will compensate for this by relying more heavily on other knowledge sources, e.g. contextual information. Stanovich's attempts to incorporate what is known about skilled and less skilled reading are reflected in the so-called interactive-compensatory model (Stanovich, 1980; 1984). Compensatory processing implies that "a deficit in any knowledge results in heavier reliance on other knowledge sources, regardless of their level in the processing hierarchy" (Stanovich, 1980, p.63). For example, a poor reader who is inaccurate and slow in word recognition but knows much about the topic of the text, will rely more on his or her background knowledge, whereas a reader with strong word recognition skills who comes upon an unknown word in the text may compensate for this lack of knowledge by analysing the form of the unknown word and by using sentence-context information. Moreover, the model explains that also skilled readers will necessarily have to rely more on higher level processing to recognize words if they read a text containing words they are not familiar with, e.g. a text for specialized purposes.

This means that an interactive-compensatory model can explain the reading process of any individual reader in any individual reading activity. Therefore, the model is a valuable contribution to the development of reading theory (cf. Samuels & Kamil, 1984/1988).

The three types of interactive reading models discussed in this section - schema-theoretic, text-based and compensatory models - have to be considered as complementary rather than competing views on reading. It is, nowadays, widely accepted that, during reading, prior knowledge directs processes of text-inferencing and global comprehension, that strong word recognition skills are advantageous for

reading comprehension and that a reader can compensate for knowledge sources that are (temporarily) weak by relying on better developed knowledge sources.

2.3. Summary and conclusion

In this chapter current theories on human information processing and reading comprehension were reviewed. Reading is considered, nowadays, to be an interactive process, that is, an interplay between lower level decoding and higher level interpretative processing. Three types of interactive models were discussed: schema-theoretic, text-based and interactive-compensatory views. Schema-theoretic views on reading explain how prior knowledge facilitates reading comprehension, whereas text-based views focus on the importance of well-developed word recognition skills. Finally, interactive-compensatory views explain that deficits in any knowledge source result in heavier reliance on other knowledge sources. Readers may compensate for weak word recognition skills by relying more on prior knowledge and vice versa. Examples of each of these three models of interactive reading were presented in this chapter. These three views together form a satisfactory picture of how interactive reading works.

The models discussed in this chapter were developed from the perspective of first language research. The next chapter will discuss whether these models also explain non-native reading, i.e. foreign and second language reading, or whether a different model is needed.

Chapter 3

Non-native reading and the use of context

The present chapter focuses on non-native reading. The point of view adopted in this study is that native and non-native reading are not fundamentally different. In both cases it is an interactive process: an interplay of lower level and higher level processing. There is, however, a difference between native and non-native reading considering the contribution of lower and higher level processing to reading comprehension. As outlined in the previous chapter, proficient native readers will possess the necessary linguistic knowledge and skills to rapidly recognize words. Non-native readers, on the contrary, hardly ever will have a native-like level of linguistic competence in the target language. This means that they will necessarily have to rely more on less efficient processing, such as conscious decoding and predicting what a text is about. In our view, it is essential that the teaching of reading in the foreign language classroom includes the training of reading strategies (including compensatory strategies). Therefore, the second part of this chapter is about 'context use'. A definition is given of what is meant by 'context use' in this study and an explanation as to why we believe that context-use training is a relevant factor in the teaching of foreign language reading comprehension.

Section 3.1 is a review of non-native reading research results. In section 3.2 variables of non-native reading are discussed. In 3.3 Bernhardt's model of non-native reading is outlined. Section 3.4 is about the use of context. Finally, in section 3.5, a summary of the present chapter is given.

3.1. Non-native reading: Review of research results

Influenced by first language reading research of the role of content schemata, early non-native reading research concentrates on the importance of knowledge-based processing. More recent non-native reading research focuses on the role of text-based processing. This last branch of research has put forward the so-called 'threshold hypothesis', that is, the assumption of the existence of a critical level of non-native proficiency needed to successfully apply native language reading strategies.

3.1.1. Knowledge-based processing and non-native reading

Earlier non-native reading research underlines the importance of a reader's prior knowledge of the topic of a text. The conclusion of these studies is that a text with a familiar text topic will be better understood than a text with an unfamiliar topic and,

moreover, that relevant background knowledge can be activated.

Omaggio (1979) found that visual stimuli (images, pictures) activated his students' content-schemata. American students reading French understood a text better after having seen an image referring to the beginning of the text story.

Adams (1982) investigated the effects of so-called 'script activators': statements telling what a text is about. The results of his research showed that script activators had a significant effect on the recognition of (nonsense) target words, both in first and foreign language reading.

Johnson (1982) showed that a text on a familiar topic is better recalled by non-native readers than a similar text on an unfamiliar topic.

Hudson (1982/1988) elaborated systematic 'pre-reading' exercises to activate helpful schemata, entailing looking at a set of cue pictures, answering questions and writing personal predictions about text contents. The results of his study showed that systematic schema-activation facilitated comprehension at the beginning and intermediate level of second language proficiency, but was of less importance for comprehension at the advanced level.

Alderson & Urquhart (1985/1988) found evidence for the hypothesis that students of non-native English belonging to a particular academic discipline were better able to understand English texts about discipline-related topics than students from other disciplines.

A study carried out by Levine & Haus (1985) demonstrated that students with 'high' prior knowledge performed significantly better in questions about text contents than students with 'low' prior knowledge. They concluded that at a high foreign language proficiency level background knowledge even had more effect on comprehension than at a lower level.

Following a study reported by Carrell in 1983, Lee (1986) distinguished three components of background knowledge: 1. prior background knowledge about text content (*familiar* vs. *novel*); 2. prior knowledge that a text is about a particular content area (*context* vs. *no-context*); and 3. the amount of lexical items in the text that actually reveal the text's content area (*transparent* vs. *opaque*). Lee's results indicated that, in a complex way, each of these three components of background knowledge played a role in non-native reading comprehension. For example, the presence of context (a title page and picture) in relation to a familiar text topic (Washing Clothes) facilitated text comprehension significantly, but comprehension of a text with a non-familiar topic (Balloon Serenade) appeared to be more successful without 'help' from a title and picture. Lee (1986) explained this as being caused by the novelty of both picture and topic. "Reading the passage does not help clarify the picture, the picture does not help clarify the passage" (p.353).

A specific type of prior knowledge is *culture-related* information. Topics that are strongly related to the culture of native speakers cause comprehension problems for non-native readers when the information in the text is not part of their own cultural background knowledge (cf. Saville-Troike, 1979; Steffensen & Joag-Dev, 1984). There are several studies in which the relevance of cultural background has been demonstrated.

Carrell (1981) found a strong relationship between the subjects' cultural background and their ability to understand folk-tales.

Johnson (1981) demonstrated that, in the case of non-native readers, text comprehension depended more on the cultural origin of a story than on the level of syntactic and semantic complexity.

Floyd & Carrell (1987) proved that training of cultural schemata positively influenced the recall of a culture-related text and the performance on questions about this text, whereas differences in the syntactic level of the text (syntactically easier versus syntactically more difficult) did not influence reading comprehension.

3.1.2. Text-based processing and non-native reading

More recent research is concerned with the role of linguistic proficiency in non-native reading. The overall conclusion of these studies is that problems of non-native readers are mainly due to deficits in non-native language proficiency. Research results show that beginning non-native reading contrasts with both skilled native and advanced non-native reading, whereas advanced non-native reading resembles the reading of a skilled first language reader (cf. Barnett, 1989).

The results of a study reported by Cziko (1980) showed remarkable differences between the reading of beginning and advanced non-native readers. Lower-level students of French heavily relied on graphic information and hardly took into account possible contextual information; advanced readers of non-native French and native French readers, however, used the same type of interactive strategies (graphic and contextual information).

Bernhardt (1987) revealed that low-level American readers of German read slowly, tended to fixate on words and only reached about 30% comprehension, while native German and advanced American readers of German only fixated on words briefly, read quickly and comprehended about 80 % of text contents.

Sarig (1987) did not find significant differences between strategy use and level of comprehension of native readers of English and advanced non-native Hebrew readers of English.

Haynes (1993) found that the ability to guess word meanings (of nonsense words) was limited by the ability to comprehend the words around the nonsense words. Readers with a low non-native proficiency level were less successful at guessing word meanings because they had less vocabulary.

Mandell (1993) examined the extent to which success in word recognition is related to successful reading comprehension in native (English) and non-native (Spanish) reading. Mandell concentrated on word-recognition processes that were not automatized, i.e. decoding activities. He used two versions of an experimental text, in which nine nouns were replaced by artificial words that conformed to the general orthographic conventions of English (in the English text) or Spanish (in the Spanish text). Success in word recognition was measured by means of the students' definitions of the artificial words. The results of his study indicated that the native readers were more successful in word recognition and reading comprehension than the non-native readers. However, there was no significant relationship between successful word recognition (decoding skills) and comprehension in either language.

Both in the case of the native and non-native readers, successful reading comprehension was related to a reduced reliance on individual word-meaning associations. Mandell explained these findings by the interactive nature of information processing. Experienced readers exhibit automaticity in the processing of words. As (native or non-native) readers decrease their focus on the structure of individual words in a text, other information sources will be accessed, thereby increasing comprehension.

Wolf (1993) investigated the relationship between word class and word meaning inferencing strategies in native and non-native reading (Spanish). Two texts were used and seven or eight words of these texts were replaced by nonsense words, created according to Spanish phonological and morphological rules. The results of her study demonstrated that at an introductory stage of non-native reading students found it difficult to deduce meaning and identify grammatical class. They tended to rely more on a combination of contextual clues and knowledge of syntax. Intermediate students relied on a variety of information sources, such as contextual clues, morphological and syntactic information. In addition, their ability to recognize morphological and syntactic cues was far more automatized than that of the lower-level students. Advanced non-native students relied on a wide array of contextual, morphological and syntactic cues, rarely misguessed the meaning of words, nor misidentified word class. They often used more contextual information and background knowledge than the beginning and intermediate students. Moreover, their ability to guess words and identify grammatical class did not significantly differ from the performance of native speakers on the same tasks. Finally, the native readers most frequently relied on contextual information and background knowledge. Processing of morphological and syntactic cues appeared to be automatic.

3.1.3. Language threshold hypothesis

The recognition of the relevance of linguistic competence has led to the hypothesis that there exists a *language threshold* (Clarke 1980/1988; Alderson, 1984), that is, a critical level of non-native language proficiency required for a successful transfer of skilled native reading strategies. According to this hypothesis, poor non-native reading is the result of reading strategies in the first language not being transferred to the non-native language, due to inadequate non-native language skills. Once non-native readers possess a certain amount of non-native proficiency, non-native reading will essentially be the same process as native reading. The behaviour of skilled native readers with poor non-native language ability, however, will resemble that of poor native readers: they will use less efficient strategies and will presumably not be able to fully understand a text. Clarke (1980/1988) refers to this as 'short-circuits'.

Considering what was outlined above, Alderson (1984) put the question whether the success of non-native reading is more related to the level of non-native proficiency or to poor first language reading behaviour. In other words, is reading in a non-native language a *reading* or a *language* problem and, in case of the latter, can the existence of a language threshold be empirically demonstrated? Bossers

(1991) reviewed three studies that investigated these issues, i.e. Hacquebord (1989), Carrell (1991) and Bossers (1989, 1992). These studies confirmed that poor foreign language reading can be due to both deficits in knowledge of the non-native language and to poor first language reading skills. For beginning non-native readers, however, the relative importance of target language knowledge is more substantial than for advanced non-native readers. These results may be an indication of the existence of a language threshold, but it could not be ruled out that there are competing explanations. Bossers commented that, if a threshold exists, it can probably not be defined in absolute terms. Following Cummins (1979) and Alderson (1984), he believes that the threshold may be different for each reading task (the more demanding the task, the higher the threshold) or depends on a reader's fluency (the more fluent, the lower the threshold).

There are also researchers that do not support the threshold hypothesis. Lee (1991), for example, does not believe that low language proficiency necessarily inhibits interactive reading. He postulates that little difference exists between native and non-native reading, which he both describes as bi-orientated processes, i.e. orientated "neither from the bottom up nor from the top down" (p.200), and that even beginning readers process interactively. According to Lee, in both native and non-native reading approximately the same proportion of bottom-up and top-down processing will take place. Low linguistic proficiency can be compensated for by using non-linguistic knowledge, especially background knowledge about the topic of the text.

Lee's ideas are comparable to what is stated by Hudson in the study reviewed above. Hudson recognizes the importance of linguistic proficiency and the existence of 'short-circuits', but believes that it is sometimes possible to override the effects of low non-native proficiency by activating appropriate content schemata. This was confirmed by Hudson's research results (Hudson, 1982/1988).

3.1.4. Conclusions research review

The studies reviewed above are very diverse: they represent different periods of non-native reading research, try to answer different kinds of research questions, do not always deal with comparable levels of non-native language proficiency, do not involve the same number of subjects and use different methods and/or measuring instruments (e.g. recall protocols, questions about text contents, thinking aloud tasks). Two conclusions can be drawn.

Firstly, the research review demonstrates that there is not one single critical factor that explains successful non-native reading. There are at least four factors that may negatively affect the success of non-native reading, that is, lack of appropriate background knowledge, low non-native language proficiency, lack of effective reading strategies and the inability to transfer effective first language reading skills to non-native reading. To what degree each of these factors contribute to non-native reading comprehension is not clear. The studies outlined above show that there is not a conclusive answer to this question.

Secondly, the question whether or not there is a critical level of non-native proficiency needed to successfully apply reading strategies cannot be answered conclusively. Assuming such a threshold level exists, it will be very difficult to define what is required to reach this level.

3.2. Variables of non-native reading

In the previous chapter, first language reading was defined as an interactive process. The view adopted in this study is that non-native reading is not essentially different from native reading. Both native and non-native reading include word recognition, higher level interpretive processing and the ability to compensate for deficiencies in knowledge. The most obvious difference between native and non-native reading is that non-native readers are not likely to have the same level of language proficiency as native readers. Additionally, there may be differences in background knowledge, especially in the case of culture-related topics. If non-native readers do not possess sufficient linguistic or background knowledge, they will necessarily have to rely on other skills. Finally, reading researchers do not exclude the possibility that characteristics of the language itself influence processing. For example, Bernhardt (1987) found that (native and non-native) readers dedicated relatively more processing time to function words in German texts than to function words in English texts.

Non-native reading behaviour is often compared to poor first language reading (cf. Barnett, 1989). Non-native readers either tend to build hypotheses about text contents or tend to analyse each word and syntactic structure of a sentence. Carrell (1988) points to these situations as 'overreliance on knowledge-based processing' or '*knowledge-biased processing*', and 'overreliance on text-based processing' or '*text-biased processing*' (p.102). Compared to skilled native reading both processes are less efficient and less effective. Carrell (1988) underlines that these poor reading behaviours may be attributed to a multiplicity of causes across individuals, such as the absence of relevant schemata, the inability to activate schemata, skill deficiencies (including deficiencies in reading skills, as well as linguistic deficiencies), misconceptions about reading (e.g. knowledge-based processing is not an appropriate activity in reading) and cognitive style. Some important variables for the success of non-native reading are discussed below.

3.2.1. Linguistic knowledge

Linguistic knowledge is knowledge of the graphic symbols in a text. A distinction can be made between knowledge of letters, words, sentences and paragraphs.

The printed *letters* are the smallest linguistic elements of a text. Sometimes non-native readers have to learn a 'new' graphic system, such as in the case of speakers of English that learn how to read Russian or Chinese. In this situation comprehension will be difficult, because a reader has to go to a lot of trouble and dedicate much processing time just to be able to simply decode the language. In less

extreme cases, non-native readers have to learn a few unknown graphemes, e.g. the ñ, ÿ and j of the Spanish graphic system.

Knowledge of *words* includes knowledge of how words are formed (morphology) and knowledge of the meaning of words (vocabulary/lexicon). Several researchers believe that vocabulary knowledge is critical, both in native and in non-native reading (cf. Yoro, 1971; Stanovich, 1986; Nagy, 1988; Nation & Coady, 1988; Grabe, 1991). Non-native readers, especially beginners, will often have limited knowledge of the target language vocabulary. Research on the relationship between vocabulary knowledge and reading comprehension, however, is relatively slight (cf. Barnett, 1989).

Knowledge of *sentences* embraces knowledge of the structure of a sentence, of the function of words in a sentence and of the meaning of parts of a sentence. Sometimes non-native syntactic rules can be different from first language rules. For example, Spanish adjectives are generally placed after the corresponding noun, while in English, German or Dutch they are placed before the noun. The role of syntactic knowledge in reading is less clear than that of vocabulary (cf. Grabe, 1991). According to several researchers, syntactic knowledge is important in non-native reading, especially at a beginning stage (cf. Barnett, 1986; Devine, 1988a; Eskey, 1988; Swaffar et al., 1991).

Knowledge of *paragraphs* is defined as knowledge of the meaning and function of cohesion devices. This includes reference (words referring to others: e.g., pronouns, demonstrative adjectives), repetition (the word itself repeated, often with a definite rather than indefinite article: e.g. 'I just saw a snake!' 'Where's the snake you saw?'); substitution (often referred to as repetition without the repeated word: e.g. 'He wanted to see her but he didn't'); ellipsis (also called substitution by zero: e.g. 'I like Rick's house but I don't like Harry's'), conjunctions (cohesive devices because of their meaning rather than through reference to other parts of the text: e.g. and, but, therefore) (Halliday & Hassan, 1976; cf. Barnett, 1989). The impact of knowledge of cohesion devices is not yet clear (cf. Barnett, 1989), but several researchers assume that this type of knowledge contributes to non-native reading comprehension (cf. Williams, 1983; Van Esch, 1987; Barnett, 1989).

3.2.2. Prior content knowledge

Because non-native readers will often have to compensate for deficiencies in their non-native language proficiency, prior content knowledge can help them to understand what a text in a foreign language is about. Therefore, the relative contribution of this variable to reading comprehension will often, inevitably, be more substantial in non-native reading than in native reading.

Prior content knowledge refers to a wide array of experiences faced in life and is also called world knowledge, pragmatic knowledge, prior knowledge or background knowledge. It includes general education, knowledge of social events and world politics and knowledge of every day occurrences, such as the fact that stamps are sold in a post office and not in a butcher's and that cars usually run on petrol and not on water. Also included is knowledge of different cultures (e.g. the course of a

traditional wedding ceremony in an Arabic country, of a bull-fight in Spain and of the Dutch St. Nicholas celebration) and of the order and logical coherence of actions or occurrences. Finally, in the context of reading comprehension, it may refer to general knowledge about languages, such as the fact that words form sentences, that most sentences consist of a subject, a direct object and a predicate and that an adjective gives extra information about a noun.

In 2.2. we saw that prior content knowledge plays a central role in schema-theoretic views on reading, that it is represented in memory at various levels of specificity and that it includes both memory of loose concepts and knowledge about their relationships (Anderson & Pearson, 1984/1988). A distinction is made between three levels of information: the conceptual, episodic and planning level. These levels are connected by means of propositions (cf. 2.2.2), which identify concepts semantically, connect them with their distinctive features, relate them to personal experiences and include them in a plan, routine or sequence (cf. Boekaerts & Simons, 1993).

At the conceptual level the meaning of concepts are stored (semantic knowledge). Concepts refer to general information that can be categorized and clearly defined. For example, the concept 'square' belongs to the category 'quadrangle' and has four equal sides (definition). Concepts can only be defined by means of other concepts (e.g. a bird is an animal; a bird has wings; wings are feathered limbs that a bird uses to fly, etc.). Propositions at this conceptual level allot a concept to a certain category within a hierarchical structure of concepts (*X is a ...*) or define a concept by means of its distinctive characteristics (*X has ..., ..., etc.*). An example of such a hierarchical structure is the BIRD schema of Collins & Quillian (1969), discussed in 2.1.1.

The information at the episodic level is subjective. It is the result of the personal experiences or episodes of individuals and, consequently, different for every single person (e.g. 'My aunt Mary has a canary'). Propositions at the episodic level represent connections between personal experiences, valuations and more general information at the conceptual and planning level. For example, 'My aunt Mary is my favourite aunt'; and 'I do not like canaries, I prefer gold-fish'; or 'Scientists are inspiring people' and 'I think Amsterdam is the most exciting city in the world'.

At the planning level, sequences of actions, rules, routines and programmes are stored. This knowledge of plans connects knowledge at the conceptual and episodic level with detailed algorithmic or heuristic procedures. An algorithmic procedure is a set of fixed actions that, step by step, lead to the solution of a problem, e.g. the solving of a mathematical problem, a detailed plan of how to install the new version of Word Perfect or the recipe of how to prepare a nice Spanish 'paella'. A heuristic procedure is also a set of actions, but in this case the order is of less importance and the outcome is not necessarily problem-solving, e.g. the actions involving the planning, organizing and making of a trip or ten rules of thumb to find out why your

car does not start. An often-cited example is the restaurant scenario of Schank & Abelson (1977). This schema includes the sequence of actions one can undertake to obtain a meal in a restaurant (cf. 2.2.1). With regard to reading, often the relevance of knowledge of formal text schemata is recognized, e.g. knowledge of the rhetorical structures of a fairy tale (normally including a chronological sequence and a happy end), the development of a simple story (including a setting, a beginning, a development and an ending) and the organization of scientific articles containing the research questions, the method, the results and discussion (cf. Carrell & Eisterhold, 1983/1988).

3.2.3. Strategy use

Reading strategies may be consciously used techniques or plans to achieve comprehension, but may also refer to unconscious processes applied automatically (cf. Barnett, 1989). Strategy use in non-native reading will often be less efficient than in first language reading. Non-native readers will not be able to quickly recognize words and will not be familiar with all the words in a text. Therefore, they will have to use time-consuming compensatory strategies at the word level and rely on their background knowledge. As outlined in Chapter 2 and also in 3.2.1, linguistic skills - especially at the word level - are considered to be essential for reading comprehension. Lack of linguistic skills often limits a reader's choice of strategy use because he or she necessarily relies more on compensatory strategies.

Reading strategies can be cognitive as well as metacognitive. Strategies at a cognitive level are directly related to a specific reading task and reading goal and can differ from task to task. Examples of cognitive strategies are automatic word recognition, analysing word form, contextually guessing and predicting text contents on the basis of prior content knowledge. Strategies at a metacognitive level can be applied to any reading activity and may involve thinking about the reading process, the planning of the process, self-monitoring while reading and self-evaluation after reading (cf. Brown & Palinscar, 1982). The variable metacognition will be discussed in 3.2.3.

Several researchers have proposed inventories of reading strategies, mostly in the context of non-native reading (Hosenfeld et al., 1981; Block, 1986; Sarig, 1987; Carrell, 1989). Grabe (1993) distinguishes between five groups of strategies that contribute to reading comprehension.

1. Local strategies for improved comprehension.
2. Strategies for main idea comprehension.
3. Consciousness raising strategies.
4. Monitoring strategies.
5. Strategies that repair mis-comprehension.

Local strategies are related to attempts to comprehend the linguistic elements of a text. Examples of local strategies are skipping words, analysing word parts, using contextual clues, considering possible cognates, re-reading words or phrases,

skimming ahead for information, scanning for specific words or information which might help, using cohesion markers for connecting information and considering the organization of the text.

Strategies for main idea comprehension include techniques to establish what a text in broad outline is about. Examples of strategies for main idea comprehension are paraphrasing text information, generating questions about text contents, examining the headings and sub-headings, making notes, inferring connections between parts of the text, establishing the purpose of the text, relating textual information to personal experience and/or background, using visual information of diagrams, charts and other images, analysing the basic structure of the text, skimming ahead and translating text parts into native language.

Consciousness raising strategies embrace processes such as becoming aware of personal motivation, ability and/or needs to read a certain text. Examples of such strategies are thinking about whether the text topic links up with personal interests, relating potential information to personal experiences, establishing a personal reading goal, considering the task being assigned with the text, considering previous success with similar tasks, predicting information from the title and generating questions to anticipate or guide reading.

Monitoring strategies are related to a reader's self-guidance during a certain reading task. Examples are predicting information to follow in the text, recognizing comprehension problems, recognizing the attitude and intention of the writer and seeking outside assistance and/or group support.

Strategies that repair mis-comprehension include considering alternative responses to comprehension problems, slowing down the reading-speed, recognizing the lack of background knowledge, re-reading text parts, analysing the immediate text part and reading ahead but remembering the problem.

It is important to be aware of the fact that it is not correct to make a distinction between successful and unsuccessful reading strategies. Whether a strategy is successful and leads to comprehension, mostly depends on other factors than the strategy itself. For example, the analysis of the form of an unknown word may contribute to comprehension in the case of a reader who only needs to identify a few unknown words of a text, but may impede comprehension if a reader needs to decode every single word. It is possible, however, to distinguish between efficient and less efficient strategies. For example, automatic word recognition is always a faster process than consciously decoding a word or extracting its meaning on the basis of contextual information and prior content knowledge.

3.2.4. Metacognition

In 3.2, it was outlined that non-native reading is often compared with poor first language reading. Non-native readers tend to overrely on textual information or on their background knowledge (cf. Carrell, 1988). To avoid this poor reading behaviour, it is necessary that non-native readers know how to approach a text in a foreign language and how to monitor their reading. This means that they will have to become aware of metacognitive aspects of reading.

Flavell (1976) defines 'metacognition' as the knowledge of one's own cognitive ability and the cognitive ability of others. Nowadays, the term 'metacognition' is used in a less strict interpretation, including a person's ability to make judgments and predictions about his cognitive functioning and to direct his own cognitive activities (cf Brown, 1980). Metacognition includes metacognitive knowledge (or 'knowing that') and metacognitive skills ('knowing how')

Metacognitive knowledge means knowledge of cognitive processes or activities, such as thinking, remembering and reasoning. Metacognitive knowledge about reading may include knowing that it is sometimes useful to reread a text part, that a word can have different meanings, that contextual information can be useful to extract the meaning of unknown words and that there is a difference between technical reading (decoding, phonological coding) and reading comprehension (understanding the message) (cf Boekaerts & Simons, 1993).

To apply or use metacognitive knowledge it has to be 'translated' into concrete activities, by connecting it with procedural knowledge ('knowing how'). These procedures or metacognitive skills (cf. Boekaerts & Simons, 1993) are the decisions that individuals take before, during and after a cognitive task (reading, writing, thinking, etc.). This entails activities of orientating oneself towards a certain task, planning and monitoring the process, diagnosing the cause of problems, evaluating what has been learnt and reflecting on the process as a whole (cf. Vermunt, 1992). In addition, goal-setting forms part of the metacognitive skills. With regard to reading, goal-setting refers to both the ability to set a goal for a specific reading task (including the intention of reading and the preferred level of understanding) and the ability to reach this intended goal by monitoring the reading process (cf. Bernhardt, 1991). For example, reading an article to reproduce the information at an examination demands a different plan or approach than skimming the headlines of a newspaper to form a global impression of today's news.

3.2.5. Other variables

Besides variables that directly relate to aspects of reading itself, other variables may affect success in native and non-native reading comprehension.

Some researchers distinguish different types of non-native readers, i.e. variables related to a reader's personal and individual preference of approaching a text. Hosenfeld (1977), for example, observed differences between readers that decode every single word and readers that sometimes skip unknown or apparently unimportant words in order to understand a text. Devine (1988b) observed comparable reading behaviour and distinguishes between sound-, word- and meaning-centered readers.

In addition to a reader's cognitive activities, his or her emotional and physical condition plays a role. The reader can be tired, distracted or feel just the opposite. These affective variables include the reader's attitude, motivation, concentration and interest with regard to a specific reading-activity. Affective variables relate to activities of dealing with feelings that, during the processing of information, lead to

an emotional state of mind that may - positively or negatively - affect the processing. Examples of such activities are self-motivation, concentration, self-estimation, doing one's best, giving subjective values to reading-tasks and dealing with emotions (Vermunt, 1992). LeLoup (1993) is an example of a study that concentrated on the affective variable personal interest in text topic. LeLoup found that high interest facilitated the comprehension of non-native (Spanish) texts.

3.3. A model of non-native reading

In the previous chapter, we concluded that schema-theoretic, text-based and compensatory models of first language reading have to be considered as complementary rather than competing. Bernhardt (1991) makes a comparable observation about non-native reading research. She refers to the history of non-native reading research as "driven by a search for the 'smoking gun'" (p.32). Instead of concentrating on the process of non-native reading as a complex of factors, research often concentrates on one individual aspect of reading (e.g. background knowledge, linguistic proficiency, first language reading ability) claiming that this particular variable explains reading comprehension. According to Bernhardt this is not a productive, nor an appropriate approach. She, therefore, suggests a holistic whole-process view on non-native reading. In such an approach, linguistic competence is considered to be essential, but not sufficient, because it will always be mediated by other factors, such as prior background knowledge and (cognitive and metacognitive) reading skills. Or rather, in the words of Bernhardt (1991), *world knowledge*, i.e. "... the intrapersonal knowledge that readers bring with them when they interact with texts" (p.37) and *literacy*, i.e. "... operational knowledge that refers to knowing how to approach a text, why one approaches it and what to do with it when a text is approached." (p.35).

These insights are reflected in a model of non-native reading (Bernhardt, 1991). This model explains how language components, literacy components and world knowledge components interact at different levels of non-native proficiency. The model explains that beginning non-native reading is characterized by a substantial amount of word-recognition errors, especially if non-native readers have to learn a totally 'new' graphic system, such as in the case of speakers of English who learn how to read Russian or Chinese. Word-recognition errors will decrease with the increase of non-native proficiency in general and non-native vocabulary knowledge in particular. The processing of syntax features shows a completely different development over time. The amount of syntax errors is relatively low at the beginning of non-native reading, increases as a result of greater exposure to and development in the language and again decreases at a relatively high proficiency level. Monitoring ability, embracing literacy and the use of background knowledge, is reader-dependent and not necessarily related to non-native proficiency or amount of exposure to the language. According to Bernhardt's model, literacy and world knowledge are relevant factors in both advanced and beginning non-native reading and highly related to strategies used in first language reading. Native and non-native

reading are more likely to overlap with regard to the use of world knowledge and literacy skills than concerning linguistic proficiency.

As Bernhardt states, "The linguistic task of all readers is to make word recognition, lexical and syntactic systems work automatically and simultaneously" (p.35). However, most non-native readers, especially beginners, will not have native-like control over the target language linguistic systems and will rely more on conscious decoding and prior content knowledge. The more literate, the higher the probability that first language reading strategies are transferred to non-native reading.

3.4. Context use

Knowledge of the nature of reading, such as discussed so far, is indispensable for theory building. One has to be careful, however, to directly apply insights into the nature of reading to the practice of teaching reading in the classroom. There is hardly ever a direct link between theory and teaching practice.

The studies outlined in this chapter reflect that non-native reading is an interactive process, requiring, besides linguistic proficiency, higher level cognitive and metacognitive skills. This implies that students will have to work on expansion of vocabulary as well as of metacognitive insights. The question, however, is how this can best be done in the teaching practice of non-native reading.

In the view adopted in this study, it is important that beginning readers in a foreign language learn how to approach a foreign language text. Due to lack of foreign language proficiency they will often show poor reading behaviour such as overreliance on text-based or knowledge-based processing (cf. Carrell, 1988). It is important, therefore, that they learn how to effectively use the knowledge and skills they possess. In this study, it is postulated that *context-use training* helps students to improve their reading skills. The term 'context' refers to the contextual information used while reading. Context-use training embraces the teaching of a set of reading strategies to help students to deal with foreign language texts and to infer the meaning of unknown words in a text. A distinction is made between *general* and *local* context. General context means extra-textual information, whereas local context is the graphic information of a text.

Based on a study by Wildman & Kling (1979), Van Esch (1987) defines the different types of context as follows:

1. General context: the reader's knowledge of the topic of a text.
2. Local context: the words and sentences around an unknown word in a text. This local context can be subdivided into:
 - a. *Syntactic context*: the information that is provided by the syntactic relationships between the unknown word and other words in the same sentence or larger text parts.
 - b. *Semantic context*: the information provided by the semantic relationships

between the unknown word and other words in the same sentence or larger text parts.

- c. *Lexical context*: the information provided by the form of the unknown word or of its composing parts.
- d. *Stylistic context*: the information that text style provides for the choice of the correct meaning of the unknown word.

The distinction between the different types of context is a distinction at an operational level. Most readers will not use the different types of context one at a time, but in an interactive way. Whether there will be much or little context available for the guessing of unknown words during a concrete reading task depends on both the individual reader's linguistic and non-linguistic knowledge and his or her ability to apply this knowledge by means of appropriate strategies.

The question whether the use of context is a relevant variable for foreign language reading will be explored by means of a field-experiment. The methodology of the experiment and the contents of context-use training will be extensively discussed in the following chapter.

Before outlining the research methodology, some examples will be given of earlier reading research that explored the effects of the training of reading strategies.

Westhoff (1981) established the effects of an intensive training programme in the use of prior knowledge during reading. He divided this knowledge into five fields: 1. knowledge of letter-combinations; 2. knowledge of sentence-patterns; 3. knowledge of semantically related combinations; 4. knowledge of logical structures (including the meaning and function of logical connectors, e.g. 'however', 'but', 'moreover'); 5. knowledge of the world. The students in his study were exposed to a training programme during an entire school year. This programme focused on the use of knowledge from fields four and five, but also included the teaching of how to use knowledge of the other three fields. The control group followed an intensive training programme in solving text-related multiple-choice questions. The subjects were native Dutch students, reading in French or German. The results demonstrated that the experimental treatment was successful. The students that had followed the training programme in the use of reading strategies scored significantly better on a cloze test than the control group. However, they did not show significant better results on a text with multiple-choice questions.

Van Esch (1987) investigated whether the use of context played a significant role in foreign language reading. The results of his study showed significant positive effects of a training programme in the use of context for beginning learners of Spanish. Because the current work is a follow-up of Van Esch's study, the latter will be comprehensively discussed in the next chapter.

Kern (1989) taught his students specific reading strategies in order to be able to guess the meaning of unknown words and understand larger text parts. These strategies are 1. recognizing cognates, prefixes, suffixes and orthographic cues; 2. recognizing and utilizing sentence and discourse level signalling cues (e.g., if ...

then, relative clauses); 3. recognizing cohesion markers (e.g., therefore); 4. inferring the meaning of unfamiliar words; and 5. finding main ideas. The control group in this experiment followed a regular reading instruction. Kern gathered three types of data: self-reports about strategy use and level of comprehension (interviews), results of a reading comprehension and word inferencing test and, finally, questionnaires about the students' language background and their own perception of their reading behaviour in non-native French. The results of the text comprehension and word inferencing tests showed significant positive effects of the training programme in the use of context. Moreover, it appeared that weak non-native readers benefited most from the teaching programme and that combinations of strategies led to better comprehension than strategies used in isolation.

Mulder (1996) compared the effects of five 'methods' to deal with teaching reading in the classroom: 1. no training; 2. traditional 'common practice'; 3. blind training (i.e. training reading comprehension ability, in which metacognitive learning activities are *not* explicitly elicited); 4. heuristic rules (i.e. metacognitive learning activities consisting of pre-given, ready-made plans); 5. open reflection (i.e. making students aware of plans to solve reading problems). These five training programmes were executed, dispersed over an entire school-year. The outcomes of this study showed that the training of reading (strategies) tended to lead to better results than no training at all. However, the differences found, were not always statistically significant. Mulder concluded that 'open reflection' leads to the best results in reading comprehension. In this type of reading instruction students themselves choose what strategies they use and plan their own metacognitive processing. An interesting finding of Mulder's study was that also students of the control condition, labeled 'common practice', substantially improved their reading comprehension ability.

Recently, the effectiveness of strategy training has been object of criticism. Influenced by text-based approaches towards non-native reading, including the 'language-threshold' hypothesis, the training of reading strategies has been considered to be beside the point. Instead, some researchers argue, it will be better to dedicate attention to the expansion of the students' foreign language vocabulary, the development of their decoding skills and other linguistic aspects of reading comprehension. Once the students reach the critical level of foreign language proficiency, they will be able to transfer their first language reading skills to a non-native reading setting. The discussion between Bossers and Westhoff is an example of these different views on the teaching of non-native reading (cf. Bossers, 1993; cf. Westhoff, 1993).

In this study, the training of reading strategies is considered to be important, also in a non-native setting. According to Stanovich's theory set out in Chapter 2 and Bernhardt's model outlined in this chapter, strategy use and metacognitive skills are as important as linguistic proficiency, both in native (cf. Stanovich, 1984; 1991) and non-native reading (cf. Bernhardt, 1991). Following Hudson (1982/1988) and Lee (1991), we believe that it is possible to overrule low non-native proficiency by

means of appropriate strategies. Therefore, a foreign language reading course should include the training of strategy use, as well as the development of linguistic proficiency. Training word-recognition skills is probably only necessary for poor non-native readers and readers that have to learn a different graphic system, such as for Russian or Chinese (cf. Bernhardt, 1991).

Using 'context' plays a role in both native and non-native reading, but is probably more important in the case of the latter. It is very likely that non-native readers need higher-level strategies to overrule deficits in linguistic proficiency. The most obvious solution is using a dictionary, but searching for every unknown word of a text takes a lot of time and interrupts reading. In this study, it is argued that a better strategy will be to try to contextually guess the meaning of unknown words and use the dictionary only if it is really necessary. This will save time, make the reader less dependent from outside help and will increase the joy of reading.

Additionally, we assume that guessing contextually is important in a more indirect way. Firstly, it will stimulate efficient dictionary use. If a reader eventually has to make use of a dictionary but has previously tried to solve comprehension problems by means of context use, he or she will already have a global idea about the meaning and word class of an unknown word. It will be easier, therefore, to make a choice out of the meanings the dictionary offers. Additionally, readers will possibly learn new words (Saragi et al., 1978; e.g. Schouten-van Parreren, 1985; Mondria, 1996).

A final argument in favour of strategy training is that students need guidance in their metacognitive processing. Bensoussan & Laufer (1984) find that most non-native students do not effectively use context strategies when left to themselves. Moreover, some students will be afraid to rely on their own judgments. Perkins & Brutten (1983) find that reader-specific variables can influence context use. According to their study, learners vary in their willingness to guess and take risks: some need more context to feel safe than others.

3.5. Summary and conclusion

In this chapter, different aspects of non-native reading were discussed. Non-native reading requires the same skills as native reading, but the relative contribution of these skills to reading comprehension is different. Non-native readers will often have to compensate for low non-native language proficiency. This affects the way a text is processed. In this chapter, a choice was made for an interactive approach towards non-native reading. This implies that the success of non-native reading is determined by a complex of factors. Variables of non-native reading are linguistic proficiency (especially vocabulary knowledge), prior content knowledge, metacognitive skills and the appropriate strategies to exploit knowledge and metacognitive insights in the process of reading. Bernhardt's model of non-native reading, discussed in this chapter, explains how each of these factors influences reading at different stages of non-native competence (Bernhardt, 1991). The exact

implications of these insights for the teaching practice of foreign language reading are not yet clear. In the view adopted in this study, however, it is important that foreign language readers, especially beginners, learn to make use of knowledge and skills they already possess to compensate for deficits in knowledge of the target language. In other words, they will have to learn to make use of the context. We defined context use as applying knowledge of the text topic (general context) and of the syntactic, semantic, lexical and stylistic information provided by the text itself (local context). It is expected in this study that context-use training may be useful for the teaching of non-native reading. This will be explored by means of an experiment. The methodology of this experiment will be outlined in the next chapter.

Chapter 4

Research plan and execution

In this chapter the plan and execution of our field experiment are outlined. The aim of this experiment was to assess the effects of two seven-week training programmes in context-use, i.e. a training programme for learners of Spanish at an elementary level and a training programme for learners of English at an intermediate level. We aimed at establishing whether the training programmes in context-use would be more successful than two corresponding control reading programmes, which represented a more 'traditional' approach of teaching reading comprehension in the foreign language classroom. It was, in fact, a multiple treatment, involving two experimental training programmes (a Spanish and English version) and their respective control reading programmes.

The two training programmes both included seven reading texts, instructions about how to make use of context and exercises in the use of context. They were comparable in the sense that they trained the use of the same context strategies and included the same type of exercises. The texts of the two programmes, however, had distinct characteristics. They were written in two different foreign languages and were destined for two different foreign language proficiency levels (Spanish for beginners and English at an intermediate level). Finally, the topics of the texts of the two programmes were not identical.

The control programmes incorporated the same seven reading texts as the training programmes. However, the exercises that accompanied the texts were different. The control programmes were assumed to reflect the usual way teachers deal with texts in a foreign language course. The control programmes did not include context-use exercises. Two tasks that were illustrative of the control programmes were reading aloud and translation exercises.

One year before the main experiment started, a pilot experiment was carried out. This pilot primarily served to improve the programmes and instruments for the main experiment. The results of the pilot also led to changes in the design, planning and organization of the main experiment.

The present study is the follow-up of an earlier study of the effects of context-use training (Van Esch, 1987). First of all, a review of the earlier project is presented in section 4.1. The aim of the current project is outlined in 4.2. Section 4.3 deals with the pilot experiment. Subsequently, in sections 4.4 to 4.6 inclusive, an outline of the research design, the subjects and the planning of the main experiment is given. Sections 4.7 to 4.9 inclusive are a description of the contents of the training

programmes in context-use, the control reading programmes and the instruments. Finally, in section 4.10, some aspects of the execution of the main experiment are reported.

4.1. The 1987-project

Van Esch (1987) hypothesized that in an interactive approach towards foreign language reading (cf. Chapter 3) context-use plays a significant role. He assumed that in the foreign language classroom weaker context-users need explicit training in order to learn how to use context. Based on these assumptions four research questions were formulated:

1. Is it possible to make a valid distinction between strong and weak context-users?
2. Will weak context-users after a training programme in context-use perform better on a context test?
3. Will weak context-users after a training programme in context-use perform better on a reading comprehension test?
4. Is there a relationship between four reader-dependent variables on the one hand and the actual performance on context and reading comprehension tests on the other?

The reader-dependent variables alluded to in the last-mentioned research question are 'foreign language vocabulary knowledge', 'attitude towards the foreign language', 'self-confidence' and 'the students' estimate of the degree to which they use contextual information while reading a foreign language'.

To find the answers to these questions, a group of beginners of Spanish was exposed to a training programme in context-use. Part of the students that performed relatively poorly on a pre-test measuring context-use - the so-called *context test* - were allotted to the experimental condition, the others took part in the control condition. Additionally, the students that performed relatively well were allotted to the control condition.

Both students of the experimental and control conditions followed a seven-week programme. The texts and most of the tasks of both conditions were similar. For example, both training and control groups performed tasks such as extracting the meaning of unknown words. The surplus-value of the experimental condition consisted of systematic training in context-use through supporting exercises, schemes and feedback of the teacher. Each lesson started with a pre-reading activity to mobilize the students' prior knowledge of the text-topic (general context) and the students also learnt how to use the information of the text itself (local context). The exercises and their solutions were discussed in the classroom and after each lesson an inventory was made of what had been learnt. The students of the control condition made the same tasks individually in the classroom, but without supporting

exercises, feedback or any further explanation of the teacher.

Van Esch found that the answer to the first three questions was in the affirmative. Based on the results of a context-test (pre-test), a distinction was made between weak and strong context- users before starting the experimental treatments. The results of the reading comprehension pre-test showed a comparable differentiation between these two groups. These differences were statistically significant. He concluded, therefore, that it was possible to distinguish between weak (C-) and strong (C+) context-users (first research question).

The group of weak context users was randomly subdivided into two subgroups. One of these subgroups took part in the training programme in the use of context, whereas the other subgroup was exposed to the control reading programme. The C+ students were all allotted to the control condition. The results showed that the gain scores of the C- group that had been exposed to the training programme in the use of context were significantly higher than the gain scores of the C- group of the control group. They did not only perform significantly better on the use of context but also on reading comprehension. Van Esch concluded, therefore, that the training programme in the use of context had been effective (second and third research question).

Also the fourth, more explorative, research question could be answered. 'Attitude towards Spanish', 'self-confidence' and 'the students' estimate of the extent to which they use contextual information while reading a foreign language' were measured by means of questionnaires before and after the experimental and control treatments. A factor analysis of the results of the questionnaires statistically distinguished between three factors, corresponding to the three variables mentioned above, i.e. attitude, self-confidence and the students' estimate of their context use. Van Esch found that, on the whole, high scores on these three factors combined with high scores on the context and reading comprehension tests. Finally, it could be established that individual vocabulary knowledge of Spanish was a relevant variable for both context-use and reading comprehension in this foreign language. Before the programmes started all subjects performed a vocabulary test. The performance on this test correlated significantly with the scores on the context and reading comprehension pre- and post-tests.

Although this earlier experiment demonstrated that context-use is a relevant variable of reading comprehension in a foreign language, further research was necessary to establish whether these results could be replicated.

4.2. Aim of the follow-up investigation

In the final remarks of his dissertation, Van Esch (1987) concluded that further research of the effects of the training programme in context-use would be necessary because of the following arguments. In the first place, a follow-up study was needed to confirm the positive results of a training programme in context-use for beginners

of Spanish. Secondly, it would not only be important to execute cross-validation research with respect to another, comparable group of beginners of Spanish, but also with groups of more advanced learners of Spanish or learners of other foreign languages. Finally, in a follow-up study, the problem of the validity of the reading comprehension test might be tackled.

This last argument needs some further explanation. In the 1987-project it was deliberately decided to use open cloze tests to measure reading comprehension ability because it turned out to be a very sensitive instrument that was able to establish improvements achieved in a relatively short period of time. It was not possible, however, to establish the criterion validity of the open cloze test because it is not clear what exactly is measured by means of this test. An open cloze test is a reading text in which every *n*-th word in a range of five to 15 has been deleted (see e.g. Taylor, 1953). The students have to fill in the deletions. The results of cloze tests correlate highly with the results of other reading comprehension tests (see e.g. Rankin, 1965; Ruddell, 1976; Bormuth, 1967; Potter, 1968; Jongsma, 1970; Anderson, 1976; Hinofotis, 1976; Stansfield, 1980). However, this type of testing is also suitable to measuring vocabulary knowledge or grammar (Alderson, 1979; Klein-Braley, 1981, 1982). Because no criterion-instrument for reading comprehension was used, Van Esch could not conclusively prove that what he had measured was indeed an improvement in reading and not an improvement in vocabulary knowledge or grammar.

The first preparations of the follow-up research started in 1989. The research plan was an adjusted version of the research plan of the 1987-project. The main differences between this and the earlier project are explained below.

We decided that, besides Spanish, the research would be extended to English as a foreign language.

Instead of an open cloze test, we planned to use a multiple-choice variant with deliberately chosen deletions and, moreover, samples of reading comprehension tests developed by CITO, the Dutch National Institute for Educational Measurement.

The control reading programmes were planned to be different from the control programme used in the previous experiment. As identified in the previous section, the actual tasks of the control programme of the 1987-project were comparable to those of the training programme in the use of context. It was, therefore, not unlikely that the significant positive effects of the training programme in the use of context found in the 1987-experiment (also) reflected the positive effects of the availability of explicit help from a teacher and not (only) the differences in the use of context and text comprehension. To manipulate a more apparent difference between the training and control conditions, the control programme of the follow-up would reflect the way teachers usually work with texts in the foreign language classroom. This means that the purpose of this follow-up is somewhat different from that of the earlier project. In the current project, we will establish whether a training programme in the use of context is more successful for reading comprehension than a more 'traditional' approach towards the reading of texts in the foreign language

classroom.

In the follow-up the four reader-dependent variables, defined in research question 3, would only be measured after the experimental treatments, in the post-test sessions. In the 1987-project vocabulary knowledge was measured by means of a pre-test, whereas 'attitude towards the foreign language', 'self-confidence' and 'the students' estimate of the degree to which they use contextual information while reading a foreign language' were measured both before and after the experiment.

Finally, we decided not to divide the subjects into a group of strong and weak context-users.

By means of the follow-up experiment presented in this chapter the same hypothesis was tested as in the 1987- project, that is, context-use is a relevant variable for reading comprehension in a modern foreign language. Context-use is the independent variable, reading comprehension the dependent variable.

Three research questions were formulated:

1. Does training in the use of context for beginning learners of Spanish as a foreign language positively influence both the use of context and reading comprehension while reading Spanish texts?
2. Does training in the use of context for intermediate learners of English as a foreign language have a differential effect on the learner's performance on the use of context and reading comprehension in comparison with a training programme for beginning learners of Spanish?
3. Is there a relationship between four reader-dependent variables - 'foreign language vocabulary knowledge', 'attitude towards the foreign language', 'self-confidence' and 'the students' estimate of the degree to which they use contextual information while reading a foreign language' - and the actual performance on context and reading comprehension tests?

The first and third question are directly based on the 1987- project. The second question is an extension of this earlier study. The experiment would not just be limited to beginners of Spanish, but also include intermediate learners of English.

The role of language proficiency is of considerable interest. As outlined in the previous chapter, researchers emphasize the importance of non-native language proficiency. Therefore, it is interesting to find out whether there will be differences between the effects of a context-use training programme at an elementary and at an intermediate non-native proficiency level.

4.3. Pilot experiment

One year before the main experiment took place, a pilot experiment was conducted. This pilot was executed at the same schools that would be involved in the main experiment, one year later (see also 4.5). The main goal of the pilot was to try out the programmes and tests. Secondly, we wanted to check whether the time planned for the experiment as a whole, i.e. September-December, would be sufficient. Finally, one of the goals of the pilot was to find out whether it would be useful - considering the limited time planned for the experiment and the consequences for the planning and design - to include a fourth research question as well. This fourth research question, formulated in the first version of the research plan, ran as follows: *Will there be a transfer effect of a training programme in context-use in one of the two foreign languages to the performance on context-use and reading comprehension in the other foreign language?*

The pilot showed that the experiment as a whole, such as executed during the pilot, was too time-consuming for the teachers and students involved. The teachers reported that they had to work under pressure in order to be able to finish the programmes and tests at the appointed time (before the Christmas holidays) and that this interfered with the preparations for the final examinations that would take place at the end of the school year. Due to this, most of the school classes involved in the pilot had to skip parts of the programmes and post-tests.

According to the teachers involved in the pilot, the main problem was the large number of tests. Because the groups of the experiment were not randomly selected, it was necessary to execute pre- and post-tests (see also 4.4). The pre- and post-tests both consisted of three sessions, spread over a three-week period. Considering the fact that the execution of the (training or control) programmes took about seven weeks, the teachers considered the time allotted to the making of the tests to be out of proportion.

To be able to establish a possible transfer effect (research question 4) it was necessary that the subjects participating in the pilot took part in an English and a Spanish programme. If they followed the context-use training for Spanish, they were allotted to the control programme for English and vice versa. Additionally, the control group had to perform a control reading programme for both Spanish and English. It was a prerequisite that, before the experiment started, the subjects had already been following the Spanish and English language courses. The number of students that met this criterion, however, was very limited. There were only three groups of students (school classes) that followed both the English and Spanish lessons. One of these groups was involved in the Spanish training programme, one in the English training programme and the third group performed two control programmes (one for Spanish and one for English). To raise the total number of subjects, also students were selected that only followed the lessons in one of the two foreign languages involved, i.e. Spanish or English. These students only performed a control programme. Accordingly, the design of the pilot experiment was rather

complex (cf. Appendix I).

The design of the pilot required that the subjects that followed a double programme should be tested for both English and Spanish. This implied that in each of the two test periods (pre- and post-tests) they had to be tested on two vocabulary tests, two context tests and four reading comprehension tests! This led to a loss of motivation.

Because of the low number of subjects in the training condition and the large amount of missing data, it was not possible to make a reliable comparison between training and control groups.

The tryout of programmes and tests will be extensively discussed in the following sections of this chapter: 4.7 (training programmes in the use of context), 4.8 (control reading programmes) and 4.9 (tests). A survey of the statistics of the tests of the pilot is presented in Appendix II.

Two different schools were involved in the pilot. These schools would also cooperate in the main experiment, one year later (see also 4.5). The data gathered in the pre-tests of the pilot did not give cause for suspecting that the subjects of the two schools would not be comparable: only in one occasion there was a significant difference between the two schools at the beginning of the experiment (Spanish reading test 2: texts with multiple-choice questions; $p \leq .01$) (cf. Appendix III).

Considering the problems discussed above, two main decisions were made after the pilot. The first decision was to cut down the number of test sessions in the main experiment, i.e. two sessions instead of three. The second decision was to drop the fourth research question. There were three reasons for the latter decision: to avoid loss of motivation, to simplify the research design and to increase the number of subjects in the experimental condition. In the following sections the lay-out of the main experiment is outlined.

4.4. Research design

As mentioned in the 4.2, the main experiment included two experimental and two control treatments. One of the experimental treatments focused on beginners of Spanish, the second on intermediate learners of English. The same applied for the control treatments.

In Figure 4.1 the design of the experiment is presented. It is a pre-test post-test control group design (Huck et al., 1974).

The four groups of the research design consisted of already established classes of students at two schools in Germany. These groups were not selected on a random basis and were, therefore, possibly not equal. Thus, it was necessary to include pre-tests in order to establish the level of each group.

We opted for maintaining the existing division in school classes. As explained above, in the 1987-project the students were first divided into weak and strong context-users on the basis of test results. Only a subgroup of weak context-users (C-) followed a context-use training programme. Such a categorization immediately at the beginning of an experiment may lead to loss of information about the differences between subjects and, consequently, to a less sensitive analysis. Other motives were to deviate as little as possible from the regular practice at the two schools (ecological validity) and to avoid bias caused by the fact that students would perform differently, once they knew that they were involved in an experiment. This is the so-called *Hawthorne Effect* (Huck et al., 1974, p.265). Finally, it was not possible to carry out the experiment using random groups for practical reasons in that this would disrupt the schools' timetables. Therefore, the existing classes were maintained.

		Pre-test				Post-test	
		S	E	S	E	S	E
School 1	Group 1	O		T		O	
	Group 2		O		T		O
School 2	Group 3	O		C		O	
	Group 4		O		C		O

S = Spanish

E = English

T = Training Programme in Context-use

C = Control Reading Programme

O = Observation (pre- or post-test)

Figure 4.1. Research design

The groups of the research design all consisted of two or more school classes. Group 1 incorporated two classes that followed the training programme in context-use for Spanish (TS1 and TS2), Group 2 three classes that followed a comparable training programme for English (TE1, TE2 and TE3). Group 3 and 4 served as control groups for respectively Spanish and English. These groups both consisted of two school classes (CS1 and CS2 for Spanish; CE1 and CE2 for English). This resulted in a total of nine school classes participating in the experiment. Appendix IV illustrates the way in which the four groups were subdivided into classes.

4.5. Subjects

The subjects were 153 students of the 'Berufsbildende Schule' (Vocational Education) in Germany. They were all native speakers of German. The age of the students was 18-21 years. The subjects all belonged to the sub-department 'Hohere Handelsschule' (Economic Administrative Education), abbreviated HHS. They followed a two-year vocational training programme that primarily concentrated on business and economics. In addition, the subjects followed one or two foreign language courses. Each subject had to follow the English course and could also choose between French, Spanish or Dutch as a second foreign language. The time planned for a foreign language course was three lessons per week (45 minutes), over two school years. During these foreign language lessons, attention was paid to commercial correspondence, language skills (especially conversation and writing), grammar and vocabulary.

Two schools were involved in the experiment, one in Kleve and a second one in Geldern. The experiment included English and Spanish as foreign languages. Each of the subjects had already studied English at secondary school over a period of five or six years and during the first year of their vocational training in Kleve or Geldern. Spanish, however, was relatively new to them. They started studying Spanish in their first year of vocational training and the experiment took place in the second and last year of the Economic Administrative Education.

The school in Kleve had also been involved in the 1987-project. One of the Spanish teachers was already familiar with the training programme and was considered to be an experienced trainer. Additionally, the experiment was a follow-up to the 1987-project and it seemed only logical to approach the same school.

There were two reasons for approaching a second school. The first reason was to limit bias. It was likely that some of the teachers, even if they had not already participated in the 1987-project, would already be familiar with the objectives and contents of the training programme in the use of context. We, therefore, decided *not* to allot classes of this school to the control condition. A comparable school was approached for the lessons of the control reading programmes: the 'Berufsbildende Schule' in Geldern, a slightly smaller town, 35 kilometres from Kleve. The teachers of this school knew that the school in Kleve participated in the project, but they were unaware of the contents and objectives of the training programmes that would be carried out in Kleve. They performed the seven lessons of the control reading programmes, which included the same texts as the training programmes, but without the exercises and information about context-use. They were told that the aim of the control programmes was to improve the foreign language reading skills of their students.

The second reason to approach the school in Geldern was a simple and practical one. Involving a second school meant that a larger number of students could participate in the experiment.

During the experiment, 77 of the 153 subjects followed a training or control programme for English and 76 subjects followed a training or control programme for Spanish. A group of 74 subjects was exposed to a training programme in the use of context, of which 33 followed the Spanish variant and 41 the English one. The other subjects, a total of 79, were exposed to a control reading programme. From this group, 43 subjects followed the control reading programme for English and 36 the control programme for Spanish. The number of subjects per class and the gender of the subjects can be seen in Appendix V.

Seven different teachers, divided between the two schools, participated in the experiment: three teachers of Spanish and four of English. In Kleve, each class had a different teacher. Only two teachers participated in Geldern, one for the two Spanish control classes, another for the two English classes. Two teachers were female and five were male.

4.6. Planning of the experiment

In the 1987-project it had become clear that the time planned for the lessons of the training programme, 45 minutes per lesson, was often not sufficient. The students' low proficiency level of Spanish appeared to be the main problem. It took them too long to read the texts and, therefore, little time was left to perform the tasks. Therefore, we decided that the follow-up would take place one semester later, in the first half of the second school year (September-December). At that time, the subjects had been studying Spanish for one year and English for approximately six or seven years. We assumed that the students' foreign language proficiency level would be sufficient to follow the training programme without problems.

The experiment was planned to last throughout the first semester of the school year. However, before the Summer holiday, during May and June, the subjects already practised the three test types that would be used during the experiment: a context test, a multiple-choice cloze test and a text with multiple-choice questions. Each test took about 30 minutes. During this preparatory stage, special attention was paid to the multiple-choice tasks, because the German students were unfamiliar with multiple-choice tests.

The *pre-test* sessions took place in September, in the second and third week of the new school year. They included a vocabulary test (tryout), a context test and a reading comprehension test. Because the pilot had shown that three test sessions would take too much time (cf. 4.3), only two sessions were organized to perform these tests. In the first session the students performed the vocabulary and context tests, in the second session the reading comprehension test. The sessions were planned to take about one hour each.

The seven lessons of the *training and control reading programmes* took place during the two following months. Each lesson was planned to take 45 minutes. At the end of the seventh lesson of the programmes, the questionnaire 'Attitude' had to be filled in by the subjects.

The *post-test* sessions took place in the second and third week of December. Besides a vocabulary test, a context test and a reading comprehension test, the post-tests involved a set of questionnaires to measure the students' level of self-confidence and their estimate of the extent to which they use contextual information. Just like the pre-tests, the post-tests consisted of two one-hour sessions.

The time between pre-test and post-test was approximately three months. The experiment was planned to be concluded just before the Christmas holidays. In Appendix VI a timetable of the planning stages of the experiment is presented.

4.7. Experimental condition

Two training programmes were developed to manipulate the independent variable context use: a programme for Spanish and a programme for English. These two programmes included the same type of exercises, but other texts. The texts of the two programmes linked up with the proficiency level of the subjects (beginners of Spanish and intermediate students of English).

4.7.1. Spanish training programme

The training programme for beginners of Spanish was based on the programme used in the 1987-project. However, some adjustments had to be made.

In the 1987-project 45 minutes per lesson was often not sufficient to perform all the exercises of a particular lesson. After discussing this problem with the teachers of the training classes, two types of exercises were excluded: *Explain the meaning of the punctuation marks in the texts* and *Explain the meaning of demonstrative pronouns*. We decided that these exercises might be excluded because other tasks of the programme which were preserved, included making use of punctuation marks and demonstrative pronouns.

Two texts of the programme were replaced by other texts. The text *El windsurf* was replaced by *El Metro de Madrid*, because in the evaluation of the training programme in 1987 the students complained that they had found the windsurf text boring. The text *El Metro de Madrid*, a text about the subway in Madrid, which had been used in the preparatory reading programme of the 1987-project, was considered to be interesting. The second text, *Extranjeros en España*, a text about foreigners living in Spain, was outdated and replaced by a more current text about Spaniards who leave their home villages to go and find a job in the big cities: *Del campo a la ciudad*.

In the 1987-project, the texts of the preparatory reading programme and training or control programme had already been lexically and grammatically analysed and adjusted to the language proficiency level of the subjects. The topics of the texts fitted into the Spanish curriculum and learning objectives, because they related to cultural knowledge. The 'new' text of the programme, i.e. *Del campo a la ciudad*, was also analysed and, where necessary, adjusted.

An important exercise of the training programme in the use of context included

the inferring of the meaning of unknown words considered to be important for text comprehension. It was essential that the students were not familiar with the meaning of these words. In the 1987-project such words had been identified by carrying out an examination of the students' textbook. A follow-up check was necessary because the research began one semester later than in the earlier project. This analysis was carried out as follows. If words did not appear in the lessons of the textbooks used in Kleve and Geldern, then they were considered to be unfamiliar. In Kleve, the first 20 lessons of a textbook titled *¡Eso sí! I* (Håkanson et al., 1990) were checked and in Geldern the first 33 lessons of the textbook *El español de hoy* (Brügel, 1989). If one of the words that had to be guessed appeared in one of the lessons of the textbooks, then it was replaced by a synonym that did not appear in the word lists of the textbooks. This was the same procedure that had been used in the 1987-project.

Although the texts of the programmes would probably not contain unknown grammatical structures - the experiment took place at a later stage of the Spanish language course than in the 1987-project - also a grammatical analysis was carried out. A list of grammatical topics was made that had been dealt with in the textbooks of Kleve and Geldern at the beginning of the experiment. This was then compared with the grammatical structures used in the texts of our programmes. As expected, it was not necessary to make adjustments.

The Spanish training programme consisted of seven texts with supporting exercises in the use of context. The shortest text contained 302 words, the longest 421. In order to avoid ambiguities about the understanding of the intention of the exercises all exercises were formulated in German, the native language of the subjects.

For each lesson written instructions were provided for the teachers of the training programme. An explanation was given as to whether the exercises had to be made individually, in groups or in couples and how long each exercise had to last. In addition, the key to the exercises was given, along with some possible ways to achieve the right answers.

The execution of the Spanish training programme in the pilot experiment was satisfactory. The teacher that executed the programme had also been involved in the earlier project and was, therefore, already familiar with the context-use training. However, we were also confronted with some problems. Classroom observations revealed that the students had to get used to dealing with the tasks of the programme and the teacher reported that the students had not always been pleased with the lessons of the programme. At first, they were excited about the 'new' method. After a few lessons, however, they complained that the way of dealing with texts was very different from what they were used to and that the exercises in the use of context were very difficult. They believed that it would have been more efficient if the teacher just had given them the meaning of the unknown words, instead of letting them perform the context exercises. More than once, the teacher explained the goal of the training programme, but he did not completely succeed in convincing the students. The dissatisfaction of the students was strengthened by the fact that they

also had to perform a control reading programme for English, which did not include context-use exercises (cf. Appendix I).

After the pilot, some measures were taken (cf. 4.3) We expected that the decision that the subjects of the main experiment would follow only one programme (a training or a control programme) should - at least partly - solve the problems described above We decided not to make any alterations in the contents of the training programme for Spanish.

4.7.2. English training programme

The training programme for intermediate students of English was developed especially for this follow-up project. The English training programme was comparable to the Spanish one in the sense that it also consisted of seven reading texts and contained the same types of exercises to teach the use of context.

Because it was difficult to establish the exact level of the subjects' proficiency of English, the teachers were asked to assist. They were asked to select the seven grammatically and lexically most suitable texts from a sample of about 20. This selection of texts was based on the teachers' experiences with these and other groups of comparable students and their personal intuitions about their students' proficiency level. It was not based on a systematic analysis of vocabulary and grammatical issues, such as in the case of Spanish. It was difficult to establish valid criteria to make such an analysis. There was no vocabulary list of the words all students were assumed to be familiar with and important grammatical issues had already been dealt with in the classroom. The relatively advanced students had different backgrounds: they had been to different secondary schools, had learnt English from different textbooks and had listened to English pop music. During the six or seven years they had been studying English, they had had a lot of opportunities to be exposed to the English language. Therefore, the teachers' intuitions were used as a criterion. The teachers not only helped to select the texts of the training programme, but also to choose the words of which the meaning had to be contextually inferred.

The topics of the texts matched the learning objectives such as formulated in the English curriculum. The shortest text of the English programme contained 376 words, the largest 417. The English programme also included written instructions for the teachers.

The English training programme was tested in the pilot experiment. The teacher that executed the training programme was satisfied and did not experience serious problems during the execution. It was the first time that he was confronted with context-use training and he was enthusiastic about the lessons. The written instructions were a necessary and helpful tool to organize and conduct the lessons of the programme such as planned. According to the teacher, it was not necessary to adjust the texts and exercises. Classroom observations revealed that the students were able to cope with the context-use exercises and enjoyed working with the programme, although at the end of the programme they were somewhat less motivated than at the beginning. The teacher believed that this was due to the fact that they got used to the tasks of the training programme.

There were no reasons to alter the English training programme such as executed in the pilot experiment. We, therefore, decided to use this programme in the main experiment without any adaptations.

4.7.3. Justification of the contents of the training programmes

In Appendix VII, the contents of the Spanish and English version of the training programme are presented. The survey shows which exercises the students had to perform, the time planned for each task and the procedure adopted (individual, couples or class).

In Chapter 3, foreign language reading was defined as an interactive process, in which both lower level and higher level processing play a role. Additionally, it was argued that it is likely that non-native readers need to be trained in strategy use to compensate for low language proficiency. Such training should not just include cognitive aspects of reading, but also metacognitive skills. This view is reflected in the training programmes in the use of context. One of the purposes of the training programmes is to teach students to use context in order to establish the meaning of important words in a text that are unfamiliar to them. They learn how they can guess the meaning of these words on the ground of what they already know of the text topic and the preceding text parts (top-down processing). Additionally, the use of context serves to develop bottom-up skills. For example, the training programmes explain how words can be formed (prefix-stem-suffix) and how the identification of the class of an unknown word can be of use in the process of establishing its meaning. In the training programmes attention is paid to each type of context use defined and explained in the previous chapter, except for use of the stylistic context. Information about the latter will only be helpful at a foreign language proficiency level much higher than the level of the subjects involved in our project. We expect that by means of the seven lessons of the training programmes, the students will learn how to use the information of the general and local context, which, in its turn, is expected to lead to both an improvement in context use and in reading comprehension.

Each of the seven lessons of the training programmes is structured in a comparable way and involves three or four exercises in the use of context. Three levels of context use are distinguished: the macro-level (general context), the meso-level (general and local context) and micro-level (local context).

The first exercise of each lesson is always an exercise that practises the use of the general context (macro-level). It is a so-called 'pre-reading activity'. The text topic is introduced and some general questions about text contents are asked. It is expected that these questions will stimulate the students to remember what they already know about the topic of the text.

In the second exercise attention is drawn to both general and local context. The exercise trains the identification and understanding of the main points of the text (meso-level). This exercise includes choosing subtitles for the text paragraphs, answering true/false-questions about text contents or predicting text contents on the basis of the subtitles of the different text fragments.

With a third exercise special attention is drawn to the use of the local context. This exercise involves inferring the correct meaning of some important but still unknown words in the text (micro-level). For this exercise the students have to work with a survey with five context strategies (or *context rules*). These five strategies refer to the types of local context, such as distinguished in the previous chapter (cf. 3.4). The survey includes the five following strategies or 'rules' (cf. Van Esch, 1987):

1. To establish the meaning of the sentence or text part in which the unknown word occurs (semantic context).
2. To establish the word class and word composition of the unknown word (lexical context).
3. To look for relationships between the unknown word and other words or groups of words in the text (syntactic context and/or semantic context).
4. To establish if the unknown word has a positive, negative or neutral meaning (lexical, syntactic and semantic context).
5. To predict the meaning of the unknown word and check whether the word class and meaning fit in the text (lexical, syntactic and semantic context).

In the course of the seven-week programme, the students are stimulated to perform the exercises without the help of the survey with context rules in order to learn to use context independently and, moreover, to transfer what they have learned to tasks not directly related to the training programme.

Finally, the training programmes in the use of context pay attention to the meaning and function of logical connectors. In this study, it is postulated that knowledge of logical connectors will facilitate text processing, including contextually guessing (cf. 3.2.1). As an introduction to the exercises that practise the use of logical connectors, in the second and third lesson of our training programme a survey of the meaning and function of some important logical connectors is presented. A distinction is made between four types of functions:

1. Analogy.
2. Contradistinction.
3. Cause and consequence.
4. Time.

In the survey, examples are given of each of the four categories. For example, the Spanish *sin embargo* means 'however' or 'nevertheless' and implies a contrast or contradistinction (function).

The first category of logical connectors (function: *analogy*) includes, among others, 'and', 'or', 'moreover' and 'in addition'. The second category (function: *contradistinction*) contains 'but', 'nevertheless' and 'although'. Examples of the third category are 'for', 'because' (function: *cause*), 'so' and 'therefore' (function: *consequence*). The fourth category (function: *time*) comprises 'before', 'after' and

'while'.

The training programmes involve two types of exercises with logical connectors. The first type includes multiple-choice exercises in which students have to choose the right logical connector out of three possible solutions and motivate their choice. The second type are exercises in which students recognize and underline logical connectors in a text, give their meaning and function and illustrate the function of a logical connector by means of concrete examples in the text. After the third lesson of the training programmes, the subjects perform the exercises without help of the survey, for the same reason as explained above, i.e. to encourage self-reliant working. Practising the use and function of logical connectors stimulates local context use, especially the syntactic and semantic context, and is assumed to contribute to non-native reading comprehension (cf. 3.2.1).

An important aspect of the training programmes in the use of context is repetition. Each type of exercise is repeated at least once. This gives the students the opportunity to get accustomed to the tasks, enables them to internalize the use of context and facilitates transfer of context-use skills to other situations or tasks (cf. Van Esch, 1987).

In Appendix VIII, the texts of the first lesson of the Spanish training programme and of the second lesson of the English training programme can be found.

4.8. Control condition

The control reading programmes included the same texts as the training programmes. The texts, however, were presented with different exercises. The most important ingredients of the reading programmes for control groups were reading texts parts aloud (pronunciation), answering questions about text contents and translating text parts.

The control reading programmes were assumed to reflect how teachers usually work with texts in the foreign language classroom. During the pilot experiment in 1990 the texts were given to the teachers without any exercises at all. The teachers of the subjects allotted to the control condition had been instructed to deal with the texts in the classroom just as they would normally do. The only restriction was that they had not been allowed to spend more than 40 minutes on each of the seven texts. The protocols made by the teachers and observations of these lessons served as the basis for the development of the tasks of the control reading programmes of the main experiment. These tasks included reading text parts aloud, answering questions and translation exercises.

During the pilot, it became clear that the lessons of the control condition sometimes included aspects that might be considered to be context use. Some of the teachers that executed the control reading programmes briefly introduced text-topic (general context) or pointed out the loan words and cognates of a reading text (local

context). A substantial difference between training and control programmes, however, was that during the lessons of the training programme context use was dealt with in a systematic way, whereas in the control programmes no such systematic training in the use of context took place.

To ensure that the lessons of the control condition of the main experiment would not comprise such aspects of the use of context, measures were taken. We decided not to include pre-reading activities, nor tasks in which students had to deal with loan words and cognates. A summary of the contents of the reading programmes for the control groups is given in Appendix IX. The texts of the first lesson of the Spanish control programme and of the second lesson of the English programme are reproduced in Appendix X.

4.9. Instruments

The instruments used in this follow-up were vocabulary, context and reading comprehension tests and a set of questionnaires. The questionnaires were essentially the same as had been used in the 1987-project. Some of the tests of the experiment had already been developed for the 1987-research project, other tests were developed for this follow-up. In the last case, we often used samples of standard reading comprehension tests developed by the staff of CITO (National Institute for Educational Measure) for the final examinations of students of secondary education in the Netherlands. With the exception of the context tests, all tests were in multiple-choice format. There were two practical reasons to opt for multiple-choice tests. Firstly, the results of a multiple-choice test can be easily and rapidly established. Secondly, the correct answers are established in advance. Most tests had been previously tried out during the pilot. A sample of the tests used in our experiment can be found in Appendix XI.

It was agreed with the staff of the schools that the results of the tests would be taken into consideration when the semester marks of the students would be established. This had not been the case during the pilot study. We expected that this measure would contribute to the reliability of the results of the tests.

In this section the choice of tests and questionnaires is explained. Additionally, the reliability of the tests is discussed.

4.9.1. Vocabulary tests

Discussing non-native reading in Chapter 3, it was argued that knowledge of foreign language vocabulary is a relevant variable for reading comprehension. By means of a vocabulary test it will be possible to establish whether there are correlations between the knowledge of words on the one hand and the use of context and reading comprehension, on the other.

Vocabulary test for Spanish

In the 1987-project, vocabulary was measured before the experiment. In the follow-up we decided to measure vocabulary after the experiment. The pre-tests also included a vocabulary test, but this was a tryout of the test that would be used in one of the post-test sessions.

The choice of the words of the Spanish vocabulary tests was based on the criterion that the words had to be familiar to the students both in Kleve and in Geldern. The objective of the vocabulary test was to establish to what degree the subjects actually knew the words that had been explained in their textbook. In all school classes that participated, the students had to study the vocabulary of the lessons of the textbook at home. A selection was made of the words that appeared in the word lists of the textbooks. The first version of the vocabulary test for Spanish, used in the pilot, was an adjusted version of the vocabulary test that had been used in the 1987-project. Because the execution of the experiment would take place somewhat later in the learning process of Spanish than in the 1987-project, the vocabulary test was adjusted to the higher level of lexical knowledge. In the pilot study in 1990, the vocabulary test for Spanish was an open test, consisting of 50 items.

A selection of incorrect answers of the subjects in this open test was used as alternatives of the multiple-choice items of the second version of the vocabulary test, which was tried out during one of the pre-test sessions of the main experiment. This multiple-choice version consisted of 50 items, with five options each. The reason to opt for five alternatives was to reduce the chance of guessing the right answer. In Appendix XII, the statistics of the tryout of the vocabulary test are presented. The results show that the Spanish vocabulary test had been rather easy for the students: the mean item difficulty, or mean p-value (proportion good answers), was .85 and the reliability coefficient (Cronbach's alpha) was .80. The items that appeared to discriminate best in the tryout were selected for the final version of the vocabulary test. Additionally, words were selected from the texts of the programmes.

Vocabulary test for English

The vocabulary test for English consisted of words that appeared in the texts that had been introduced during the first year of English at the schools in Germany. The vocabulary tests for English only had to meet the criterion that the words were familiar to the students in both Geldern and Kleve. Topics dealt with in the English classroom were *The language of the press*, *Dynamic Management*, *Successful Enterprise by Modern Marketing*, and *The Techniques of Advertising*. These topics appeared in each of the two textbooks that were used at the schools of the experiment. In Kleve the title of this textbook was *English im Berufsfeld Wirtschaft* (Jansen et al., 1986), in Geldern the textbook was called *Focus on Success* (Macfarlane, 1987).

The procedure to construct the vocabulary test for English was comparable to the procedure described above. The results of the tryout of the multiple-choice version of the vocabulary test outlined in Appendix XII show that the test had been suitable

for the subjects of the experiment (mean item difficulty: .59; reliability: .85).

4.9.2. Context tests

The context tests were partly developed for the 1987-project and partly for this follow-up. By means of a context test the ability of the students to infer the meaning of unknown words from contextual information is measured. A context test consists of two texts in each of which 10 words are underlined. The students have to infer the meaning of the underlined words and write them down on the answering form. We decided to use two texts with different topics to provide to the subjects two general contexts that could be used for comprehension.

The underlined words of each text had to meet three criteria:

1. The words had to be important for text contents.
2. The words had to be unknown to the students.
3. The words had to be 'guessable' from the context.

These criteria had also been applied in the 1987-project. The context post-tests for Spanish and English were different from the context pre-tests. This was a prerequisite, because one of the criteria was that the meaning of the words that had to be contextually inferred, had to be unknown to the subjects. All the words that had been 'guessed' correctly during a pre-test were, from that moment on, potentially familiar to the readers. Therefore, we decided to use comparable context post-tests with 20 'new' unknown words. In total, four context tests were used, two for Spanish and two for English.

Context tests for Spanish

For this follow-up project, the same tests were used as in the 1987-project. In the last-mentioned research, the texts of the context tests had already been grammatically and lexically analysed and adjusted to the language proficiency level of the subjects. It was important to verify whether the underlined words were also unknown to the subjects of the follow-up. This could simply be established by ensuring that these words did not already appear in the word lists of the students' textbook. No adjustments were required.

The context test of the pre-test consisted of the texts *El café, bebida internacional* and *Destrucción y conservación de la naturaleza*, the context post-test included the texts *Emilia, una mujer de Tenerife* and *Las Fallas*.

In the 1987-project the reliability of these tests had already been established. The subjects were a comparable group of 82 students that followed the same type of vocational training as the subjects of the follow-up. The reliability coefficient of the context test of the pre-test was .74. and of the post-test .65. This was not very high, but could be explained by the limited number of test items, i.e. 20 (cf. Van Esch, 1987). In our pilot, we found the following reliability coefficients: .73 (pre-test) and .78 (post-test). The results of the pilot also show that the mean item difficulties of

the context pre- and post-tests were comparable (.53 and .54) (cf. Appendix II, Tables 1 & 2).

Context tests for English

The context tests for English had to be developed for this follow-up study. The selection of the words that would be underlined had to meet the same criteria as mentioned above with regard to the tests for Spanish. The first criterion (the words had to be unknown) caused problems, because it was difficult to establish with certainty whether a word was unknown to all the students (cf. 4.7.2). Four texts were selected and of each of these texts 15 words that were presumably unfamiliar to the students. To establish whether these words were in fact unfamiliar, we decided to test a comparable group of students. Both schools also offered a three-year vocational training programme (subdivision *Gymnasial Oberstufe*). In the year that the experiment took place, a group of about 35 students studied English in the third year of this programme. Before the experiment started, i.e. just before the Summer holiday, this group was asked to perform a vocabulary test, including 60 words selected from the four texts referred to above. The words were presented as a multiple-choice vocabulary test, with five options to each word. It was assumed that if these advanced students of English did not know the meaning of these words, it was highly probable that the subjects of the experiment would not be familiar with these words either. The test was performed by 32 students and the mean score was 21.7 (reliability coefficient: .81). For each text of the context tests, the ten most difficult items (words) were selected. It was not only important to ensure that the words that had to be guessed were unknown, but also to check that the vocabulary needed to guess the meaning of the unknown words (the immediate and broader context) was familiar to the subjects or, at least, could easily be recognized (e.g. loan words).

The context pre-test included the texts *Irradiation? No thanks* and *Save the trees*, the post-test *British one-day trippers to France to buy beer* and *Making a market*. In the pilot the reliability of the English context tests was established. The coefficients were not very high, but comparable to what was found for Spanish in the 1987-project, i.e. reliability coefficients of .66 (pre-test) and .66 (post-test). The mean item difficulties were .56 (pre-test) and .53 (post-test) (cf. Appendix II, Tables 7 & 8).

4.9.3. Reading comprehension tests

By means of context tests we could establish whether the training programmes in the use of context influenced the variable context use itself. However, an even more important question in this study is whether ability in the use of context also influences reading comprehension.

Two instruments were used to measure reading comprehension. Firstly, samples of standard reading comprehension tests developed by CITO for the final examinations of secondary education in the Netherlands were used. Secondly, we opted for the use of multiple-choice cloze tests. This type of test is different from the

traditional open cloze test such as used in the 1987-project.

Some of the multiple-choice cloze tests were developed by the staff of CITO for a pilot experiment (Van Zuylen, 1991), whilst others had to be developed for the experiment. It was not possible to use German standard tests to measure reading comprehension, for the simple fact that such tests did not exist.

In the pilot two independent reading comprehension tests were applied. Both pre- and post-test included a test consisting of two reading texts with multiple-choice questions and a multiple-choice cloze test. The first-mentioned type of tests were samples of standard CITO tests. A complete CITO test, however, consists of a set of reading texts and 50 multiple-choice questions, whereas the tests of the experiment only involved two texts and 20 to 25 questions. The school boards had demanded that the tests sessions would last one hour, at a maximum. Therefore, it was not possible to use a test of 50 items. Also the second type of reading tests, the multiple-choice cloze test, included two texts.

After the pilot, we decided to reduce the test sessions to two, instead of three (cf. 4.3). In the pilot, two sessions had been planned to measure reading comprehension. We decided to reduce this to just one session. This, however, led to other problems. As explained in 4.2, instead of the open cloze test used in the previous project, we had planned to use both multiple-choice cloze tests and samples of reading comprehension tests developed by CITO. Moreover, the time planned to measure reading comprehension should be one hour, at a maximum.

After discussing this problem with the staff of CITO, we decided to develop 'cocktails' of texts with multiple-choice questions, including parts of standard CITO tests and multiple-choice cloze test. Each 'cocktail' consisted of three texts, one text with multiple-choice questions and two multiple-choice cloze tests. One of the advantages of the 'cocktails' was that the number of test items would be larger, about 35 items instead of 20-25. This should positively affect the reliability of the reading comprehension tests. Moreover, results of research carried out by CITO had demonstrated that the two types of reading tests both measured reading comprehension skills (Mets & Glas, 1988). Finally, we found significant correlations ($p \leq .01$) between the results of the two types of reading tests (pre-tests): .41 (Spanish) and .63 (English). The correlation between the Spanish reading tests was low, but this can probably be explained by the fact that the Spanish pre-tests had been rather easy for the students (cf. Appendix II, Tables 3 & 4). The reason to decide that the main part of the reading tests would consist of multiple-choice cloze tests was that performing a cloze test was less time-consuming than answering multiple-choice questions about text contents. The reading comprehension tests of the main experiment were presented to the students as follows: (A) first multiple-choice cloze test (B) test with multiple-choice questions (C) second multiple-choice cloze test. To avoid bias caused by the order in which the different test parts were presented to the students, two versions were made: A-B-C and C-B-A.

Because in a normal school situation students will not have to make the same test twice within a relatively short time, the reading comprehension test of the post-test was different from the reading comprehension test of the pre-test. The reading comprehension tests of pre- and post-tests did not always have the same number of items.

Reading comprehension tests for Spanish

Unfortunately, it was not possible to use existing CITO-tests for Spanish (texts with multiple-choice questions) without any adaptations. The subjects in Germany had followed Spanish for only one year and consequently their proficiency level was still low. The lowest standard level of the reading comprehension tests developed by CITO is meant for students of MAVO (junior general secondary education). The students of this type of education in the Netherlands follow a Spanish course for at least two years. We decided to use samples of tests at the MAVO-level and adapt the texts and questions. The simplified texts and questions in Spanish were corrected by a native speaker. Additionally, we decided that the questions had to be translated into German because we considered it to be very important that the students did not have difficulties understanding the questions. One of the teachers of Spanish in Kleve, a native speaker of German, translated the questions into German.

For the pilot, four texts were selected from of a set of MAVO examinations. This selection was based on the relatively high mean item difficulties, measured by CITO. For the pre-test of the pilot the following texts were selected: *Arantxa Sánchez Vicario* (mean item difficulty = .70) and *La Coca-cola cumple 100 años* (mean item difficulty = .63). For the post-test of the pilot the following texts were selected: *A niño de tres años* (mean item difficulty = .67) and *Kook Kafé* (mean item difficulty = .66).

The results of the pilot showed that the texts with multiple-choice questions had been rather easy for the students (pre-test: mean item difficulty = .83, reliability coefficient = .70; post-test: mean item difficulty = .69, reliability coefficient = .68) (cf. Appendix II, Tables 4 & 6). We, therefore, decided that in the main experiment the questions would only be presented in Spanish.

The multiple-choice cloze tests for Spanish were developed especially for this follow-up project, with collaboration from the staff of CITO and some teachers of the Department of Spanish of the University of Nijmegen.

A sample was made of appropriate texts. In some cases the texts had been adjusted to the level of the subjects. Then, an analysis of text structure was made, following a method proposed by Mets & Glas (1988). These two staff members of CITO had investigated the possibility of using multiple-choice cloze tests to measure reading comprehension. Based on Kintsch & Van Dijk (1978) they divided texts into macro- and micropropositions (cf. 2.2.2). Words representing the main arguments of the text (macropropositions) were deliberately deleted. For each of the deleted words, three options were given. Mets & Glas (1988) formulated the four following criteria to select which words in a text can be deleted:

1. The deleted words have to represent the main arguments (or macropropositions) of the text.
2. If the rest of the textual information is understood, it will be also be possible to fill in the deleted word.
3. To solve an item, more textual information is needed than the immediate context of the deletion.
4. Any overlap of items must be avoided.

Regarding the second criterion, they added two important sub-criteria. Firstly, the text has to be sufficiently redundant, also without the deleted words. Secondly, it has to be avoided that items can be solved just by means of background knowledge, that is, without the help of textual information. More or less the same criteria have to be applied to the two false options.

The cloze tests were read and commented on by a panel of five experts, consisting of teachers of Spanish and members of the Spanish department of CITO. On the ground of their comments, the tests were improved several times.

The three options of each deletion were presented in both Spanish and German. It was important to exclude the possibility that the students rejected one or more options just because they did not know what the word in question meant.

Finally, four texts were chosen to be used in the pilot experiment: *'Pizza' para todos* and *Fin de semana en Buenos Aires* were tried out in one of the pre-test sessions (mean item difficulty: .74; reliability coefficient: .75), *España, cambio de imagen* and *Autoestop compartido* in one of the post-test sessions (mean item difficulty: .61; reliability coefficient: .81) (cf. Appendix II, Tables 3 & 5). Because the pre-test had been rather easy for the students, in the main experiment the multiple-choice options of these tests would only be presented in Spanish. Only the words that did not appear in the word lists of the students' textbooks would also be given in German.

The following combination for the 'cocktails' of the main experiment were chosen. The pre-test of the main experiment included: (A) *Pizza para todos*, (B) *Arantxa Sánchez Vicario*, (C) *Fin de semana en Buenos Aires*. The post-test consisted of three other texts: (A) *Autoestop compartido: la forma más económica de viajar*, (B) *La Coca-Cola cumple 100 años*, (C) *España: cambio de imagen*.

Reading comprehension tests for English

Since the proficiency level of the subjects that participated in the English programmes was relatively high, it was possible to use tests developed by CITO. Before the pilot experiment took place, we established what type of secondary education in the Netherlands combined best with the level of reading comprehension of the subjects of the two schools in Germany. A selection was made of tests for three different proficiency levels, i.e. MAVO (junior general secondary education), HAVO (senior general secondary education) and VWO (pre-university education). These tests were presented to a group of German students comparable to the

subjects, but not involved in our experiment. The results of this small-scale study showed that the tests at the MAVO level were the most appropriate. Consequently, four texts from previous MAVO examinations were selected. For the pre-test of the pilot *Travelling selfishly* and *Bread winner* were used. The post-test included *Save or seals* and *How to sell with success*. The tryout of these tests showed that the pre-test results were satisfactory (mean item difficulty: .53; reliability coefficient: .74), but the post-test results not (mean item difficulty: .69; reliability coefficient: .50) (cf. Appendix II, Tables 10 & 12).

For the multiple-choice cloze tests for English, we planned to use materials developed by CITO. As mentioned above, at the time of our pilot project CITO was investigating the possibility of using multiple-choice cloze tests for the standard examinations of English. Permission to use these new materials in our pilot project was acquired. Before the pilot started a group of students - the same group that made the samples of MAVO, HAVO and VWO tests - was asked to make these tests. It appeared that two of the tests developed by CITO linked up with the proficiency level of the subjects of the experiment. The other tests were too difficult. These two tests were *Getaway with murder* (mean item difficulty, measured by CITO: .67) and *Scandal of a gap that lets children vanish* (mean item difficulty, measured by CITO: .60). These tests were used in the pre-test session. Additionally, we decided to use the test *On the brink of...* This test, however, was the only test that had four instead of three options for each deletion. To maintain uniformity, the alternatives were reduced to three. The fourth cloze test of the pilot had to be developed by ourselves. The same procedure was followed as explained above, with regard to the development of the Spanish cloze tests. Some of the teachers of English in Kleve commented upon the first versions. The final version was corrected by a native speaker of English. This test had no title. The two last-mentioned tests were used for the post-test session.

The results of the tryout of the multiple-choice cloze tests were more or less comparable with what was outlined above with regard to the texts with multiple-choice questions: the pre-test statistics were satisfactory (mean item difficulty: .55; reliability coefficient: .73), but this was not the case for the post-test statistics (mean item difficulty: .70; reliability coefficient: .57) (cf. Appendix II, Tables 9 & 11).

Regarding these results, the reading comprehension post-tests could not be validated. Therefore, these tests would not be used in the main experiment.

Fortunately, during and after the pilot CITO developed two other multiple-choice cloze tests that were appropriate for the proficiency level of the subjects of this study, i.e. *Runner Stephanie Douglas tells her story* (mean item difficulty, measured by CITO: .71) and *The missing budgies bureau* (mean item difficulty, measured by CITO: .66) (Van Zuylen, 1991). We decided to use these tests for our main experiment.

For English, the 'cocktail' of the pre-test of the main experiment included: (A) *Getaway with murder*, (B) *Bread winner*, (C) *Runner Stephanie Douglas tells her story*. The post-test consisted of (A) *The missing budgies bureau*, (B) *Travelling selfishly*, (C) *Scandal of a gap that lets children vanish*

4.9.4. Questionnaires

The post-tests involved the use of a set of questionnaires. The questionnaires had also been used in the 1987-project. The questionnaire *Attitude* of the 1987-project was an adapted version of a questionnaire developed by Spoelders (1978). The questionnaire *Self-confidence* was constructed by Van der Meijden (1984). Both questionnaires were translated into German. The two other questionnaires had been developed at the beginning of the 1987-project. The questionnaires of this follow-up study only included the questions that had contributed to three factors distinguished by means of a factor analysis, carried out in the 1987-project: 'Attitude towards the foreign language', 'Self-confidence' and 'The students' estimate of the degree to which they use contextual information while reading a foreign language'. From now on these three factors will be referred to *Attitude*, *Self-confidence* and *Estimate of context use*.

The questionnaires were only presented to the subjects in the main experiment of the follow-up project, not during the pilot. The questionnaires were of the five-scale Likert type and can be found in Appendix XIII.

Four questionnaires were used:

1. Questionnaire Attitude (17 questions).
2. Questionnaire Self-confidence (7 questions).
3. Questionnaire Context use (8 questions).
4. Questionnaire Reading Ability (5 questions).

The questionnaire *Attitude* consisted of 17 statements to which the students were asked to react. These statements referred to such aspects as whether or not they found the foreign language at issue (Spanish or English) an important subject, whether they studied the foreign language because it was a compulsory subject, whether the time dedicated to the foreign language at school was sufficient, whether they found it interesting to study the foreign language, whether they enjoyed studying the foreign language and whether they always made their homework.

The questionnaire *Self-confidence* was executed just before the students performed the context test. The students were asked questions about how they experienced performing such tests. They were asked whether they sometimes felt insecure (question 1), how well they usually performed such a test (question 2) and if they ever had doubts about their capability to perform well on a context test (question 3). Additionally, we asked whether they felt nervous if a text included words they did not know (question 4) and how they estimated their ability to answer questions about text contents (question 5). The last two questions dealt with their feelings at the moment that their teacher returned a marked test. They were asked if

they felt nervous in such a situation (question 6) and how they reacted if the result was worse than expected (question 7)

The questionnaire *Context use* was executed directly after the context test. This questionnaire checked whether the context strategies, such as explained in the training programmes (see 4.7.1), had in fact been applied by the students. The questionnaire asked what actions the students had undertaken to infer the meaning of unknown words in the text. They were asked subsequently whether they used contextual information such as prior content knowledge (question 1), illustrations (question 2), syntax and meaning of the sentence in which the unknown word was found (question 3 and 4) and the word class of the unknown word (question 5). Additionally, they were asked if they had checked whether this meaning fitted into the text semantically (question 6), syntactically (question 7) and with regard to word class (question 8).

The fourth questionnaire dealt with the students' self-confidence concerning reading comprehension tasks. This questionnaire was performed just before the reading comprehension post-test. The students had to estimate their reading comprehension ability in general (question 1 and 5) and, more specifically, their reading-speed (question 2), their ability to understand the main points of a text (question 3) and their ability to distinguish between essentials and side-issues (question 4). This and the second questionnaire together included the questions that had contributed to the factor Self-confidence.

By means of these questionnaires the assumed correlation between the variables (factors) *Attitude*, *Self-confidence* and *Estimate of context use*, such as defined above, and the actual performance of the students on the context and reading comprehension tests was verified.

We deliberately opted for executing the questionnaires only during the post-tests. The questionnaire *Context use* explicitly mentioned, one by one, the context strategies taught by means of the training programmes in the use of context. The intention was to avoid contamination between training and control conditions. If the questionnaires had been performed by the students before the experiment (pre-tests), there was a chance that also the students of the control condition would become aware of the techniques used in the training programmes.

The experiment also involved a fifth questionnaire, i.e. the students' evaluation of the project. The contents and results of this questionnaire will be discussed extensively in Chapter 6 of this study.

4.10. Execution of the main experiment

Each of the nine classes was able to finish the programme and tests in time. This means that the measures taken after the pilot experiment had been successful. However, we experienced different kinds of problems during the execution of the

main experiment.

First of all, the pre-test results of the main experiment, unexpectedly, revealed that there were huge differences between the Spanish training and control groups at the beginning of the experiment. This could implicate that the groups were not comparable. As outlined in 4.3, the results of the pilot had not given cause for suspecting that the groups for Spanish would not be comparable (cf. Appendix III). For practical reasons we decided to carry on the experiment, despite these pre-test results. Firstly, it would be not be easy to find a 'fresh' suitable population for our experiment. Secondly, the school in Kleve also had been involved in the 1987-project. Thirdly, we had already executed a pilot experiment, the programmes and tests had been adjusted to the foreign language curriculum and textbooks of the schools in Kleve and Geldern and the school boards had given their permission for the execution of the experiment. Finally, there were no reasons to believe that for English we would be confronted with the same type of problems.

Moreover, our classroom observations during the experiment revealed that there were some serious motivational problems, especially in the case of the classes involved in the Spanish programmes.

Finally, we noticed that, despite our measures (cf. 4.8), there was still some overlap between the training and control conditions for Spanish.

The execution of the main experiment will extensively be discussed in Chapter 6, which is an outline of the students' evaluation and of the teachers' opinions about the experiment. But, first of all, the results of the main experiment will be outlined in the next chapter.

Chapter 5

Research results

In the previous chapter, three research questions were formulated. The first question is whether a training programme in the use of context for beginning learners of Spanish as a foreign language positively influences both context-use and reading comprehension. The second research question is whether a training programme in context-use for intermediate learners of English as a foreign language has a differential effect on the learners' performance on context-use and reading comprehension in comparison with a training programme for beginning learners of Spanish. The third research question is whether there is a relationship between four reader-dependent variables - 'foreign language vocabulary knowledge', 'attitude towards the foreign language', 'self-confidence' and 'the student's estimate of the degree to which he or she uses contextual information while reading a foreign language' - and the actual performance on the context and reading comprehension tests of the experiment. The aim of this chapter is to answer the three research questions.

In section 5.1 the results of the context and reading comprehension tests are presented, both for Spanish and English. In section 5.2 the correlations between the subjects' scores on the four reader-dependent variables and their performance on the context and reading comprehension tests are reported. The results are summarized in section 5.3.

5.1. Results of the context and reading comprehension tests

In this section the results of the context and reading comprehension tests are reported. On the basis of these results, it can be established whether or not the training programmes in context-use, such as outlined in the previous chapter, were more successful than the control reading programmes. Concerning the elementary level (Spanish), positive effects of the training programme were expected to find because the results of a comparable experiment showed significant improvements of beginning learners of Spanish, both with regard to context-use and reading comprehension (Van Esch 1987, cf. 4.1). To find out whether, and in how far, learners at a more advanced non-native language proficiency level will be influenced by a training programme in context-use, this follow-up also involved intermediate students of English. We anticipated to find positive effects of the English training programme, but no predictions were made about whether these effects would be stronger or weaker than those of the Spanish training programme.

In the second and third chapter of this study, we saw that reading researchers are interested in the role of linguistic proficiency. It is likely that differences in non-native language proficiency influence the learning and use of reading strategies. In our case, we had the opportunity to assess the effects of a training programme in the use of context at two levels of non-native proficiency. Our results may contribute to a better understanding of the role of non-native language proficiency in the use of context and reading comprehension ability.

The overall results on the context and reading comprehension tests are reported in 5.1.1. These results do not provide clear answers to the research questions. Most of the gain scores were negative and there were reasons to suspect that the results of the Spanish groups were not comparable. A second analysis was carried out at the school class level and a third one at the individual level. The results of these two additional analyses are presented in 5.1.2 and 5.1.3.

5.1.1. Overall results

Data were gathered from two test periods, a set of pre-tests at the beginning of the experiment and a set of post-tests after the execution of the training and control treatments (cf. 4.4, Research design).

The reading comprehension tests did not always have the same number of items. Therefore, proportions (of correct answers) were used, instead of raw scores. Firstly the pre- and post-tests results were established, secondly the gain scores (score post-test minus score pre-test). At first, the best method to establish whether the differences between training and control groups were statistically significant was considered to be an analysis of covariance. However, it appeared that our data violated the underlying assumption of equality of slopes. Therefore, the independent samples t-test was used.

Results context tests

To judge the answers in the context tests (items $n=20$), the four following criteria were used (cf. Van Esch, 1987):

1. One point, for each word that was properly used in the context, both semantically and syntactically.
2. Half a point, if the word was semantically acceptable in the context, but syntactically wrong.
3. No point, if the meaning of the word did not fit the context.
4. No point, if the item had not been filled in.

For the correction of the context tests an inventory was made of accurate, half-correct and false answers. The first versions of these inventories were established before the experiment took place. In the case of Spanish, the lists of the former project were used (Van Esch, 1987). After the first correction of the context tests, the lists were updated with words that had been filled in by the subjects and did not yet appear in these inventories. In case of doubt, we asked a number of experts to

judge whether or not a solution was correct. These experts were teachers of the two schools in Germany and professors of the University of Nijmegen. Finally, a second and third correction was carried out, strictly according to the updated lists.

In Table 5.1 the results of the context tests are presented. A first inspection of the pre-test results immediately shows that there were considerable differences between the Spanish training and control groups at the beginning of the experiment. The control group scored more than twice as high on the context pre-test, leading to a difference of .40 ($p \leq .01$). Another complicating factor is that the control group's mean score on the context pre-test was very high, i.e. .78. A ceiling effect could, therefore, not be excluded.

Table 5.1. Context tests: Means and standard deviations (proportions)

	Subj. n	pre-test		post-test		gain scores	
		mean	s.d.	mean	s.d.	mean	s.d.
ST	25	.38	.23	.42	.17	.06	.14
SC	24	.78	.13	.71	.14	-.07	.13
ET	25	.52	.16	.52	.15	.00	.17
EC	18	.61	.19	.52	.21	-.09	.14

ST = Spanish Training Group - School 1

SC = Spanish Control Group - School 2

ET = English Training Group - School 1

EC = English Control Group - School 2

The overall results of the context tests show that the gain scores of both the Spanish and English training groups are (relatively) better than the gain scores of their corresponding control groups. The difference between the training and control groups for Spanish is .11. For the training and control groups for English the difference is .09. The results of the t-test are presented in Table 5.2. We had predicted that the training groups of Spanish and English would have improved their performance on the variable context use after being trained in this subskill. Therefore, one-tailed t-tests were applied. The results demonstrate that in both cases the difference between the training and control groups is significant (Spanish: $p \leq .01$; English: $p \leq .05$). It is questionable, however, whether in the case of Spanish such a comparison should be made, regarding the pre-test results. This will be discussed in Chapter 7.

Table 5.2. Context tests: Results of the t-tests

Groups	t-value	Degrees of Freedom	1-tail probability
Spanish	3.02	47	.00**
English	1.76	41	.04*

** = significant at the 1% level

* = significant at the 5% level

Results reading comprehension tests

In Table 5.3 the reading comprehension test results are presented. The pre-test results confirm that there were already differences between the Spanish training and control groups at the beginning of the experiment. The difference (.12, in favour of the control group) was not as huge as in the case of the context pre-test, but yet significant ($p \leq .01$).

Table 5.3. Reading comprehension tests: Means and standard deviations (proportions)

	Subj. n	pre-test		post-test		gain scores	
		mean	s.d.	mean	s.d.	mean	s.d.
ST	25	.53	.17	.41	.15	-.12	.24
SC	24	.65	.09	.65	.11	.00	.12
ET	25	.60	.12	.53	.12	-.07	.11
EC	18	.57	.13	.50	.14	-.07	.12

ST = Spanish Training Group - School 1

SC = Spanish Control Group - School 2

ET = English Training Group - School 1

EC = English Control Group - School 2

Unexpectedly, the training group for Spanish had a lower main gain score than its control group. The difference was .12. In the case of English, no difference between the mean gain scores of training and control groups was found. The results of two t-tests, reported in Table 5.4, show that the gain scores of the Spanish training and control groups were significantly different ($p \leq .05$).

Table 5.4. Reading comprehension tests: Results of the t-tests on the mean gain scores

Groups	t-value	Degrees of Freedom	2-tail probability
Spanish	-2.38	36.39	.02*
English	-.10	41	.92

* = significant at the 5% level

Conclusions overall results

The outcomes are different from what was found in the 1987-project. The overall results of this follow-up study do not convincingly prove that the training programme in the use of context for beginning learners of Spanish significantly improved context use and reading comprehension. There were two reasons to suspect that the results of the Spanish groups were not comparable. Firstly, already before the experimental manipulations took place, there was a considerable difference between the Spanish training and control groups. Secondly, the control group's mean score on the context pre-test was very high and a ceiling effect could not be excluded. Finally, an incompatible result was found: the Spanish training group performed significantly poorer on reading comprehension than the control group.

Concerning English, no comparison can be made with the 1987-project, because the latter only involved Spanish as a foreign language. In this study, the English training programmes positively influenced the students' performance on the context tests. However, this better performance on context use did not result in a better performance on the reading comprehension tests.

These results led to the decision to execute two additional analyses, i.e. one at the class level and an analysis at the individual level. We expected that these analyses would contribute to a better understanding of the outcomes of the experiment.

5.1.2. Results by school class

Each group of the research design consisted of two or more school classes. In this section, the results of the school classes are reviewed and compared with the overall results discussed above. In Appendix XIV the results of the Spanish and English school classes are presented.

Spanish

The results of the school classes involved in the Spanish programmes show a complex picture. In Figure 5.1 the scores of the four school classes on the context pre- and post tests are reflected. The figure demonstrates that the starting points (pre-tests results) of the four classes were not the same. The two control classes (especially CS1) performed strongly on the context pre-test, whereas the two training classes performed poorly (especially TS1). In the case of CS1, there is a ceiling effect (pre-test score of .87), meaning that the pre-test had been too easy to be able to establish possible effects of the training programme on the basis of gain scores.

TS1, with a very low pre-test score of .31, is the only school class that shows an improvement in performance on context use (context post-test). A comparison of mean gain scores of the context tests, shows that this performance was significantly better than the performance of the two control classes in Geldern ($p \leq .01$). However, we will have to be very careful to attribute this better performance to the training

Spanish: Context use

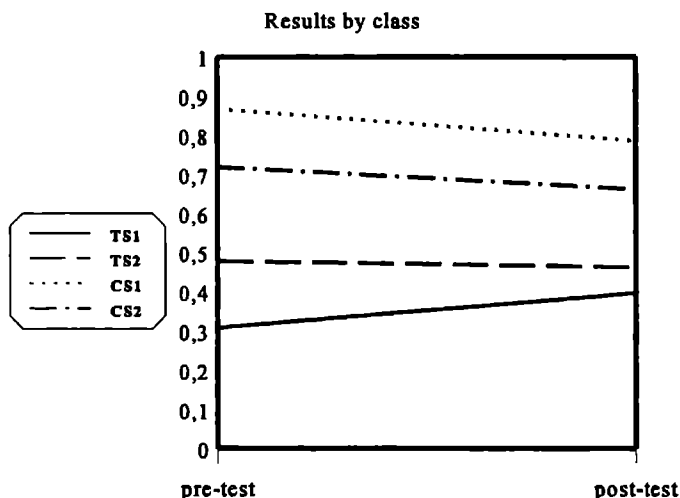


Figure 5.1.

programme in the use of context. Considering the high pre-test scores of the control classes (especially CS1) and the low pre-test score of TS1, it was more likely that the latter would score higher on the post-test than the two control classes. The mean gain score of the other school class that followed a training programme in the use of context (TS2) was not significantly different from the results of the two control classes (cf. Tables 1 & 2 of Appendix XIV).

Spanish: Reading

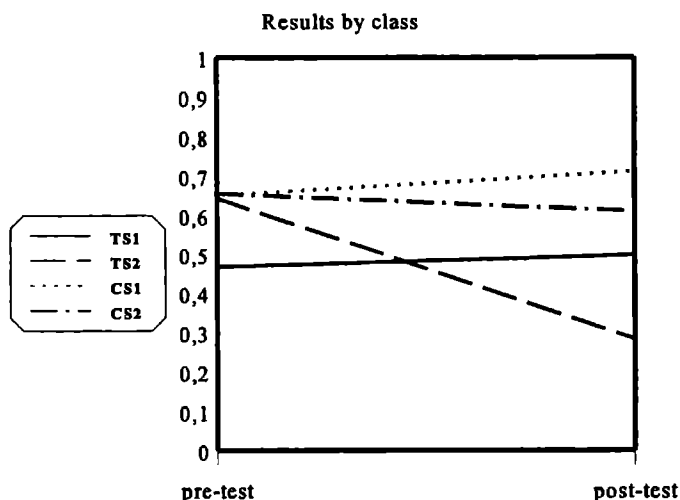


Figure 5.2.

In Figure 5.2 the scores on the reading comprehension tests of the school classes for Spanish are represented. A comparison of mean gain scores on reading comprehension demonstrates that TS1 did not perform significantly different from the two control classes. The mean gain score of TS2, on the contrary, is significantly poorer than both mean gain scores of the control classes and the mean gain score of TS1. Finally, there is a significant difference between the mean gain scores of the two control classes: the score of CS1 is significantly higher than of CS2 (cf. Tables 3 & 4 of Appendix XIV).

English

Five school classes were involved in the English lessons of our experiment. The results of four of these classes are reported in Appendix XIV. We decided not to take into account the results of TE3 because, after excluding subjects with missing data, only four subjects were left.

The results by class for English show that only one of the two training classes (TE1) performed significantly better on context use than the two control classes. The performance of the other training class (TE2) was comparable to the performance of

the two control classes (cf. Tables 5 & 6 of Appendix XIV). The performance on context use of the four school classes involved in the English part of the experiment are reflected in Figure 5.3.

English: Context use

Results by class

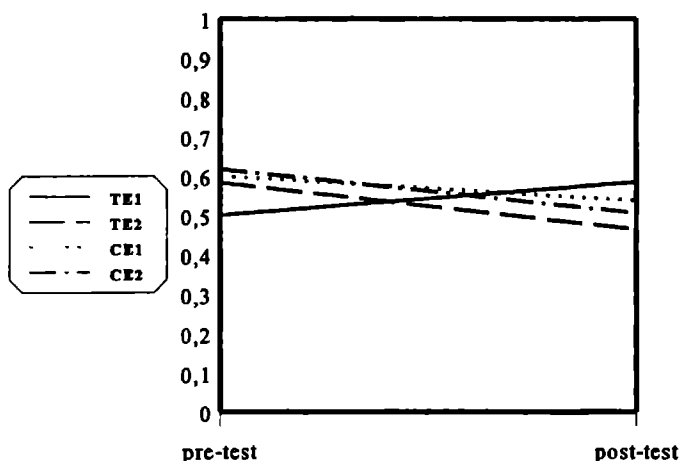


Figure 5.3.

With regard to the results of the reading comprehension tests, no significant differences between the school classes were found (cf. Tables 7 & 8 of Appendix XIV). Figure 5.4 is a diagram of the performance on reading comprehension of the four school classes.

Conclusions results by school class

The results by school class discussed in this section provide valuable additional information about the effects of the training programmes in the use of context. The analysis by school class shows that, regarding *context use*, TE1 was the only school class that performed better on the context post-test than on the context pre-test. Regarding the high pre-test results of the control classes, however, it could not be conclusively demonstrated that this better performance was indeed the result of the training programme. In the case of English, the mean gain score of TE1 was significantly better than the mean gains of the two control classes for English. The two other classes that followed a training programme, i.e. TE2 and TE3, did not show significant positive effects of the training programme with respect to their performance on the context test.

English: Reading

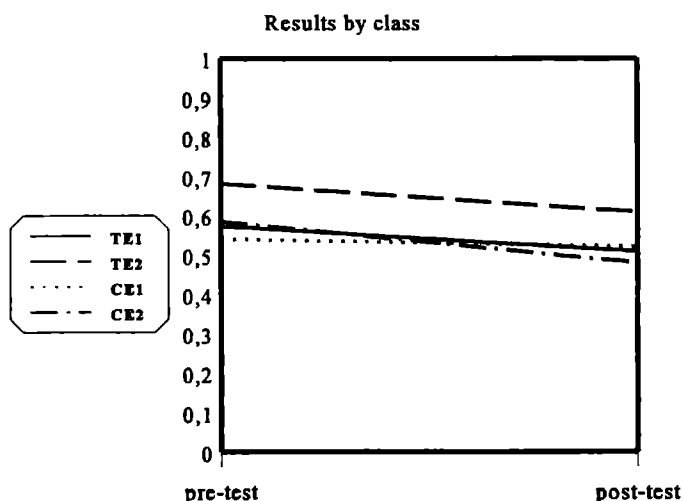


Figure 5.4.

With regard to *reading comprehension*, the results show that the performance of one single school class strongly influenced the overall results. The extremely low post-test score of one of the two classes that had followed the Spanish training (TS2) caused the significantly weaker performance of the training group for Spanish, discussed in 5.1.1. The other class that followed the training programme for Spanish (TS1) did not perform significantly better or worse than the two classes of the control condition. There was also a significant difference between the results of the two control classes: CS1 scored relatively better on reading comprehension than CS2. For English, the reading comprehension results by class confirm the overall results: no significant differences were found between the gain scores of the different school classes.

5.1.3. Results at the individual level

In Chapter 3, we saw that there are differences between the individual reader's linguistic and non-linguistic knowledge and his or her ability to use this knowledge by means of appropriate strategies. This implies that it is not unthinkable that the effects of a classroom training programme can be different for individual learners. Therefore, we decided to carry out an analysis of the performance of each individual subject. The purpose of this analysis is to establish whether individual subjects behaved differently from the overall performance of the subjects allotted to the control condition. A distinction was made between the subjects that participated for Spanish and those participating for English. Firstly, for each subject the score

(proportion of correct answers) on the pre-test was established and referred to as p_1 . Secondly, the score (proportion of correct answers) on the post-test was established, referred to as p_2 . Next, the value of p_1 and p_2 for the whole group of subjects involved in the control condition was established and the difference between these proportions (p_2 minus p_1) was calculated. This outcome was our so-called *reference-value*. For each subject, the difference between p_2 and p_1 was compared with this reference-value, by means of a test for two proportions. This implies that these two proportions were *not* compared with zero. Thus, for each subject a z-score was established that described the difference between the individual value and the reference-value. This z-score can be interpreted as a deviate of the unit normal curve (cf. Ferguson & Takane, 1989, p.199-200).

With regard to the context tests, two analyses of each context test were carried out: one so-called *flexible* analysis, by means of which all half correct answers - semantically correct, but syntactically wrong (cf. 5.1.1) - were counted as correct answers and a so-called *rigid* analysis, in which all half correct answers were counted as false answers. With regard to reading comprehension, it was not necessary to make this distinction, because the answers to the tests were either correct or wrong. The results of these analyses are presented in Table 5.5.

To limit the possibility of finding results by chance, only z-scores of individual subjects, representing a probability of one percent or (less) of finding these values, were taken into consideration. In the case of the context tests, one-tailed tests were used because of our prediction that the training programmes in the use of context would have a positive effect on performance on context use. For reading comprehension, two-tailed tests were applied. In Table 5.5, the + and - are followed by the number of subjects with extremely high (+), or extremely low (-) z-scores.

With regard to performance on *context use* the results show that, independently of the type of analysis (rigid or flexible) and of language, extremely high individual z-scores are generally found in the training groups. In the case of Spanish, the major part of the subjects that performed very well belonged to TS1. However, as outlined above (cf. 5.1.2), we cannot prove for sure that this is (only) the result of the training programme in the use of context. Considering English, the number of extremely high scores increases if the flexible analysis is applied. The major part of the subjects for English that shows extremely high z-scores belonged to TE1.

The analysis of the individual performance on *reading comprehension* shows that five subjects of the training group for Spanish scored extremely low. These subjects belonged to the same school class, i.e. TS2. The analysis at the school class level, discussed in 5.1.2, demonstrated that the gain score of this particular class was significantly poorer than that of the other school classes. The training group for English showed extreme behaviour in both positive (two subjects) and negative direction (one subject). The two subjects that scored extremely well belonged to two different school classes, i.e. TE2 and TE3. In the control groups no extreme z-scores

Table 5.5. Numbers of subjects with extremely high and low z-scores on context use and reading comprehension

	School 1 Training		School 2 Control	
Context use:	+	-	+	-
Spanish 'rigid'	3	0	1	1
Spanish 'flexible' (n=57)	3	0	1	0
English 'rigid'	3	0	0	1
English 'flexible' (n= 58)	6	0	1	1
Reading comprehension:	+	-	+	-
Spanish (n= 59)	0	5	0	0
English (n= 57)	2	1	0	0

were found.

Conclusions analysis at the individual level

The analysis at the individual level illustrates that extreme changes in individual performance in context use and reading comprehension, both in a positive and a negative sense, are almost exclusively found for subjects of the training condition. The results for context use might indicate that the training programmes positively influenced the performance of individual subjects, although in the case of Spanish there are also competing explanations. Regarding reading, for some subjects of the Spanish training group, context-use training appears to have led to extremely negative effects. The results do not give reasons to believe that there is a relationship between performance on context use and on reading comprehension: none of the individual subjects that scored extremely well on context use, showed a comparable performance on reading comprehension.

5.2. Relevance of four reader-dependent variables

In the previous section the results of the context and reading comprehension were discussed. The aim of the present section is to establish whether students' scores on four reader-dependent variables correlated with their performance on the context and reading comprehension tests of the experiment. These four variables are 'foreign language vocabulary knowledge', 'attitude towards the foreign language', 'self-confidence' and 'the student's estimate of the degree to which he or she uses contextual information while reading a foreign language'. In the previous chapter, the three last-mentioned variables were abbreviated to *Attitude*, *Self-confidence* and *Estimate of context use*. The variable 'foreign language vocabulary knowledge' will be abbreviated to *Vocabulary*.

To measure these four variables, the post-tests included a vocabulary test and questionnaires. The degree of correlation between 'vocabulary', 'attitude', 'self-confidence' and 'estimate of context use' on the one hand, and the results of the context and reading comprehension tests on the other, is established by means of the Pearson product-moment correlation coefficient.

This third research question was also explored in the earlier project (Van Esch, 1987). Van Esch found significant correlations between the vocabulary knowledge of a group of beginning students of Spanish and their performance on context use and reading comprehension. In addition, the results of the 1987-project demonstrated that, on the whole, poor context users scored lower on 'attitude', 'self-confidence' and 'estimate of context use', than strong context users (cf. 4.1). However, only part of these differences were statistically significant. Regarding the variable 'attitude', both pre- and post-test results showed a significant difference between weak and strong context users; concerning 'self-confidence', only the pre-test results. Neither pre- nor post-test results showed any significant difference between weak and strong context users with regard to the variable 'estimate of context use'. According to Van Esch, the pre-test results could be due to the lack of the students' insights into their context use at the beginning of the experiment and the post-test results to the fact that for this last analysis only the results of the control groups were considered.

Just as in the earlier project, vocabulary tests and questionnaires were used to measure the four reader-dependent variables mentioned in the third research question. There were, however, some differences between the execution of the 1987-project and the experiment reported in the current study.

In the 1987-project the correlations between 'vocabulary' and context and reading comprehension test results were established on the basis of a vocabulary test that had been performed in one of the pre-tests sessions. The pre-test of the current study also involved a vocabulary test but this test was a tryout of the test that would be used in one of the post-test sessions (cf. 4.9.1). Consequently, in this study the correlations between 'vocabulary', context use and reading comprehension are based on the performance on the vocabulary post-test.

In the earlier project both pre- and post-tests embraced questionnaires. This was not the case in the current research, in which only the post-test incorporated questionnaires (cf. 4.9.4). The consequence of this choice is that it is not possible to establish differences between performance on the questionnaires before and after the programmes, such as had been done in the 1987-project.

Finally, a different statistical technique has been applied to compare the performance on the questionnaires with the performance on context and reading comprehension tests. In the former project separate ANOVAs were executed to establish whether there were significant differences between the performance of the groups of the research design. In the current study correlation coefficients are established.

This section is organized as follows. In 5.3.1 the correlations between the students' performance on 'vocabulary' and the context and reading comprehension are presented. In 5.3.2 the correlations between 'attitude', 'self-confidence', 'estimate of context use' and performance on the context and reading comprehension tests are reported. The correlations are presented in Appendix XV.

5.2.1. Vocabulary

In Chapter 3 the importance of linguistic proficiency was underlined and especially the importance of vocabulary knowledge. It is interesting, therefore, to establish whether vocabulary knowledge played a role in this study and, if so, if vocabulary knowledge is more important at a beginning level of foreign language proficiency (Spanish) or at an intermediate level (English).

Spanish

The results for Spanish are presented in Table 1 of Appendix XV. The results are all statistically significant. The results of the vocabulary test significantly correlate with the students' performance on *context use* ($r=.64$, $p<.01$). The correlation between vocabulary knowledge and *reading comprehension*, however, is low ($r=.35$, $p=.014$).

Regarding vocabulary knowledge and context use we found a correlation coefficient that is more or less comparable with what was found in the 1987-project. In the earlier project the correlations between vocabulary knowledge, measured in one of the pre-test sessions and performance on context use were .69 (pre-test) and .51 (post-test).

Concerning the correlation between vocabulary knowledge and reading comprehension, the correlation coefficient found in our study is lower than what was found in the 1987-project. The correlation coefficients found in the earlier project were .67 (pre-test), .67 (post-test) and .83 (repeated post-test).

Our results confirm that, in the case of beginning students of Spanish, vocabulary knowledge is a relevant factor for both context use and reading comprehension. The

correlations, however, are not exceptionally high. In the case of reading comprehension, our results are different from what was found in the 1987-project.

English

In Table 2 of Appendix XV the correlation coefficients of the English vocabulary, context and reading comprehension tests are reported. The correlation between vocabulary knowledge and performance on the context test is .62 ($p \leq .01$) and between vocabulary and reading comprehension .66 ($p \leq .01$).

The results demonstrate that for English the correlation between vocabulary and *context use* is comparable with what was found for Spanish. The correlation between vocabulary and *reading comprehension* is higher than in the case of Spanish.

5.2.2. Attitude, self-confidence and estimate of context use

In this section we will establish whether there are significant correlations between the reader-dependent variables 'attitude', 'self-confidence' and 'estimate of context use' on the one hand, and performance on the context and reading comprehension post-tests on the other.

Spanish

In Table 3 of Appendix XV the correlation coefficients for Spanish are presented. The scores on *attitude* significantly correlate with the post-test results. The correlation with the context post-test is .51 ($p \leq .01$) and the correlation with the reading comprehension post-test is .32 ($p \leq .05$).

The variable *estimate of context use* does not significantly correlate with the context post-test ($r = .11$, $p = .46$), nor with the reading comprehension post-test ($r = .27$, $p = .06$).

Concerning *self-confidence*, a significant correlation is found between this variable and performance on the context post-test ($r = .28$, $p \leq .05$). The correlation coefficient, however, is low. No significant correlations are found with regard to performance on the reading comprehension post-test ($r = .03$, $p = .84$).

We found no significant correlations between the three variables that were previously distinguished, i.e. 'attitude', 'estimate of context use' and 'self-confidence'. The correlation coefficient of the variables 'attitude' and 'estimate of context use' is .11 ($p = .44$), the correlation between 'attitude' and 'self-confidence' is .06 ($p = .67$) and between 'estimate of context use' and 'self-confidence' this is .21 ($p = .14$). These results confirm the outcomes of the factor analysis carried out in the 1987-project. In other words, the results prove that, in the case of beginning students of Spanish in this and the earlier project, the variables 'attitude', 'estimate of context use' and 'self-confidence' are independent from each other.

For several reasons, these results cannot be compared with the results of the earlier project: only our post-tests included questionnaires, no comparison was made

between performance of training and control groups and a different statistical technique was used. However, we may conclude, with the necessary reservation, that both in this analysis and the analysis carried out in the 1987-project 'attitude' appears to be a relevant variable for context use and reading comprehension and that by means of both analyses is demonstrated that there are inconsistencies in the students' estimate of context use and their actual performance on the context tests.

English

In Table 4 of Appendix XV the results for English are presented. There are no significant correlations between the variable *attitude* and the test results (context test: $r=.00$, $p=.99$; reading comprehension test: $r=.29$, $p=.84$).

The variable *estimate of context use* does not significantly correlate with the context post-test ($r=.10$, $p=.52$). In addition, no significant correlations are found between this variable and the reading comprehension post-test ($r=.25$, $p=.11$).

Comparable outcomes are found with regard to the variable *self-confidence*. The results do not show any significant correlation between this variable and the test results. The correlation between 'self-confidence' and the context post-test is $-.04$ ($p=.80$) and the correlation between 'self-confidence' and the reading comprehension post-test is $.17$ ($p=.28$).

Finally, the results in Table 4 of Appendix XV demonstrate that, concerning English, no significant correlations are found between the variables 'attitude' and 'estimate of context use' ($r=.29$, $p=.08$), nor between 'estimate of context use' and 'self-confidence' ($r=.24$, $p=.12$). There is, however, a significant correlation between 'attitude' and 'self-confidence' ($r=.53$, $p=.00$). This is different from what was found for Spanish.

5.3. Summary and conclusion

In this chapter the results of our experiment were presented. On the basis of the overall results of the context and reading comprehension tests, we cannot conclude that the training programmes in the use of context were more successful than the more traditional way of dealing with texts in the control programmes. First of all, there were reasons to believe that the results of the Spanish groups were not comparable. Moreover, an unexpected result was found: the mean gain score of the Spanish training group was significantly poorer than of the control group. The training group for English scored significantly better on the context use than its control group, but no significantly positive effects of the training programme on the variable reading comprehension were found. This means that the answer to the first research question, i.e. whether a training programme in the use of context for beginning learners of Spanish as a foreign language positively influences both the use of context and reading comprehension, is negative.

The lack of overall positive results implied that it was not possible to establish whether the training programme in the use of context for beginning learners of

Spanish had been more successful than the training in the use of context for intermediate learners of English, or vice versa. Consequently, the second research question cannot be answered on the basis of the results of this study.

To gain insights into the nature of our data, the results of two additional analyses were presented: an analysis by school class and an analysis at the individual level. The analysis by school class shows that there were differences between the performance of the school classes. The analysis at the individual level demonstrates that for a limited number of students the training programme appears to have led to extremely positive effects and also extremely negative effects.

In this chapter, also the third research question was dealt with: the relevance of the four reader-dependent variables 'vocabulary', 'attitude', 'self-confidence' and 'estimate of context use'.

From the results outlined in this chapter can be concluded that, in this experiment, vocabulary knowledge significantly correlated with context use and reading comprehension, both in the case of Spanish and English. The correlation between Spanish vocabulary knowledge and reading comprehension, however, was low.

Attitude towards the foreign language correlated with the students' success in the Spanish context and reading comprehension tests. For the intermediate students of English, the variable 'attitude' did not significantly correlate with their test results. In our study, 'attitude' appears to be more relevant for beginning foreign language student than for advanced students.

The students' estimate of the degree to which they used contextual information did not combine with their actual performance on the tests of our experiment.

The students' reported level of self-confidence did not correlate with their actual test results. Only in the case of the Spanish context test a significant correlation with the variable 'self-confidence' was found. The relevance of self-confidence for context use and reading comprehension was, therefore, not clear.

A final conclusion that can be drawn on the basis of the results outlined in this chapter is that the effects of our training programmes in the use of context were dependent on a complex of factors. What factors may have played a role in the experiment will be discussed in the following chapters. In Chapter 6, the students' and teachers' evaluation of the experiment are reported. A discussion of our research results is presented in Chapter 7.

Chapter 6

Evaluation of programmes and tests

In the previous chapter the results of our experiment were reviewed. This chapter deals with the evaluations of the programmes and tests by the students and teachers that took part in the project. The students filled in a questionnaire and the teachers were asked to express their opinions in an open interview.

This chapter contains three sections. In 6.1 the results of the questionnaire are outlined and in 6.2 the results of the interviews. Section 6.3, finally, reflects how these outcomes contribute to a better understanding of the nature of this experiment and its impact on classroom practice.

6.1. The students' judgement of programmes and tests

After the experiment, the subjects filled in the five-scale Likert type questionnaire *Evaluation of Programmes and Tests* (cf. Appendix XVI). The questionnaire included 16 questions. The first nine questions dealt with the students' opinions about the (training or control) programmes. They were asked questions on aspects such as whether the lessons of the programmes had been useful, whether the aim of the programmes had been clear and whether their reading comprehension skills had improved. The questions 10 to 16 inclusive were about the tests of the experiment, i.e. whether the aim of the tests had been clear, whether there had been sufficient time to perform the tasks of the tests and whether the test results had been satisfactory. Before the actual analyses took place, the answers to question 6, 7, 8, 11, 12, 13, 14, 15 and 16 were recoded in such a way that the most positive answer was rewarded with five points and the most negative answer with one point.

The results of the questionnaire are found in Appendix XVII. The means and standard deviations of each of the four groups are presented in Table 1. For the correct interpretation of these mean scores, it is important to remember that the most negative score is 1 and the most positive score 5. In Table 2 of Appendix XVII the frequencies of negative, neutral and positive answers are given. In Tables 3, 4 & 5 the results by school class are reported.

Programmes

The students' overall appreciation of the programmes was neither high nor low. The mean score of the four research groups on questions 1 to 9 inclusive - the questions about the programmes - was between 2.73 (Spanish control group) and 3.00 (English training group).

Question 1. The English training group believed that they had *learned how to use contextual information* (3.21). The results of the frequency-analysis confirmed that 57.6% of the English training group found that they had become better context users. The Spanish training group, however, was not fully convinced that the programme had been successful (2.70). Only 20% of this group was positive about the effects of the training programme. In both cases, the control group scored slightly lower than the training group, i.e. 2.55 (Spanish) and 2.94 (English).

Question 2. The groups found the programmes neither dull nor enjoyable. The highest score was found for the English control group (2.79). The Spanish training group had the lowest score (2.27). This group included the largest percentage of students that had not liked their programme (63%). There were no substantial differences between the *pleasure* experienced by the students who followed a training programme in the use of context and those who followed a control reading programme.

Question 3. The Spanish training group estimated that they had not *improved their foreign language reading* (2.17). Only 6.7% of the students of this group believed that they had improved their ability to understand Spanish texts, whereas 56.7% estimated that they had not. The English training group was slightly more optimistic than the Spanish training group and scored 2.88. The results of the control groups, i.e. 2.32 (Spanish) and 2.47 (English), did not provide a conclusive answer to this question.

Question 4. On the whole, the *aim* of training and control programmes had been sufficiently *clear*. The lowest score was found for the Spanish control group (2.87). The mean score of the three other groups was higher than 3 points.

Question 5. The English control group found that the programme offered enough *variation* (3.15). The other groups were less positive about the diversity of the tasks resulting in the following scores: 2.77 (Spanish training group), 2.61 (Spanish control group) and 2.58 (English training group).

Question 6. Each of the four groups reported that the *results of the classroom exercises* had been more or less satisfactory. The highest score was found for the English training group (3.24). Most of the students of this group, i.e. 54.6%, had been pleased with their achievements in the classroom. The Spanish training group scored 2.90, the Spanish control group 2.85 and the English control group 3.00.

Question 7. The *amount of time assigned to the reading of the texts* of the programmes had been sufficient.

Question 8. There had been enough *time to perform the exercises* of the programmes.

Question 9. The students were neither negative nor satisfied with the *performance of their teacher* during the lessons of the programme. The English control group included the largest number of students that evaluated the teaching

negatively (50%).

The results of the analysis by school class revealed that there were sometimes different outcomes for school classes that had followed the same training or control programme (cf. Appendix XVII, Tables 3, 4, & 5).

One of the two classes of the Spanish training group, i.e. TS2, tended to evaluate the programme more negatively than TS1. For the students of TS1 the purpose of the programme had been clear (3.53) and there had been sufficient time to read the texts of the programme (3.65). The students of TS2, however, found that this was not always the case and scored respectively 2.46 and 2.61.

The results of the three school classes involved in the English training programme showed that TE2 and TE3 often scored higher than TE1. The most salient differences were found in the evaluation of the effectiveness of the programme (question 1), the pleasure experienced working with the programme (question 2) and the diversity of the programme (question 5).

Finally, there were differences between the two control classes for Spanish. The majority of CS1 was satisfied with the control reading programme, whereas a substantial part of the students of CS2 was negative about it.

Tests

In the second part of the questionnaire the tests were evaluated. The overall judgment of the tests was positive.

Question 10. The *aim* of the tests had been clear. The mean scores of the groups for Spanish, however, tended to be lower than those for English (Spanish training: 3.03; Spanish control: 2.87; English training: 3.27; English Control: 3.26).

Questions 11 to 15 inclusive. The students were satisfied with the *amount of time* that had been available to perform the tests. The English training group showed the highest results, i.e. scores between 3.12 and 3.67. The Spanish training group tended to score somewhat lower than the other groups. Their lowest score was 2.87 (question 13), their highest 3.23 (question 12). The analysis of frequencies showed that the majority of the students of the training group for Spanish was inclined to fill in the neutral answer, whereas the majority of the three other groups opted for a positive evaluation.

Question 16. The Spanish training group seemed somewhat less satisfied with the *test results* than its control group (Spanish Training: 2.80; Spanish Control: 3.19). For English, it was just the opposite (training condition: 3.27; control condition: 2.94).

Finally, the results by school class showed that TS2, one of the school classes that had followed the Spanish training programme, had not always had sufficient time to perform the tests.

In Table 6.1 the means of the total scores on the questionnaire are presented. The maximum total score was 80, the minimum total score 16. T-tests showed that there were no significant differences between the four groups (Spanish Training, Spanish Control, English Training, English Control) regarding the overall evaluation of

programmes and tests. Within two of these groups, however, there were significant differences between individual school classes. The total score of TE1 was significantly lower ($p \leq 0.01$) than the scores of TE2 and TE3 (English training group) and CS1 was significantly more positive ($p \leq 0.01$) about programmes and tests than CS2 (Spanish control group).

Table 6.1. Results questionnaire "Evaluation of Programmes and Tests". Mean total scores.

	mean	s.d	n
Spanish Training Programme	46.43	10.05	30
TS1	49.18	8.04	17
TS2	42.85	11.54	13
Spanish Control Programme	47.39	11.24	31
CS1	54.80	5.49	15
CS2	40.44	10.86	16
English Training Programme	50.88	8.66	33
TE1	43.00	8.01	12
TE2	54.44	5.70	9
TE3	56.08	4.78	12
English Control Programme	48.26	9.32	34
CE1	48.76	5.14	17
CE2	47.76	12.33	17
Entire population	48.30	9.85	128

6.2. The teachers' judgement of programmes and tests

Not only the students were asked to evaluate the programmes and tests, also the teachers were given the opportunity to express their opinions about the experiment. To gather the relevant information, the seven teachers involved in our experiment

were interviewed. These interviews had an open character. A list with key points served as point of departure. These points were 1. contents of the programmes; 2. execution of the programmes; 3. effects of training and control programmes; 4. tests; 5. questionnaires; and 6. motivation. Each interview was taped, transcribed and analysed.

6.2.1. The training and control programmes

As explained in Chapter 4, the programmes of the experiment were executed by seven different teachers. Five teachers were involved in the training programmes, one for each school class. The other two were responsible for the lessons of the control reading programmes. One of these teachers executed the control programme in the two school classes for Spanish, the other in the two classes for English. Below these teachers will be referred to as **Teacher TS1, Teacher TS2, Teacher TE1 Teacher TE2, Teacher TE3, Teacher CS1+CS2 and Teacher CE1+CE2**. Teacher TE1 and Teacher TE3 were females.

This section is divided into three parts, which refer to three aspects of the teachers' evaluation of the programmes: Contents (I), Execution (II) and Effects (III).

I. CONTENTS

Two main issues were discussed in relation with the contents of the training programmes in the use of context and the control reading programmes: the texts (topics, degree of difficulty) and the tasks (exercises in the use of context, tasks of the control programmes).

The texts of the programmes

The seven teachers (Spanish & English) who executed the training and control programmes were positive about the contents of these programmes. They considered the text topics to be interesting and to link up with the foreign language curriculum.

The two teachers of the control groups (Teacher CS1+CS2 and Teacher CE1+CE2) enjoyed working with the programme because it was a nice alternative for the materials in the textbooks normally used in their lessons. Teacher CS1+CS2 liked working with the texts of the programmes because they were far more up-to-date than the materials in their own textbook.

Favourite text topics of the Spanish training and control programmes were *ETA* (lesson 4) and *Amnesty International* (lesson 5) and of the English programmes *Swatch* (lesson 2) and *How to find a Job* (lesson 4).

The tasks of the programmes

The tasks of the training programme in the use of context were regarded to be useful and effective. Little information was provided about the tasks of the control programmes.

The two teachers of the Spanish training programme in the use of context confirmed that their students experienced many problems dealing with the tasks of

the programme.

The teachers of the training groups for English were positive about the exercises of the training programme. Their students had learnt to work independently. Also in the lessons after the training programme they tried to infer the meaning of unknown words in reading texts, instead of asking for it.

II. EXECUTION

In the interviews some aspects of the execution of the experiment were discussed, i.e., time per lesson, the teachers' time-tables, homework, the planning of the experiment, the foreign language proficiency level of the students, the students' motivation and the execution of the control programmes.

Time per lesson

The teachers had been given strict instructions about the execution of the lessons of the programmes (time-schedule). Most teachers found that 45 minutes per lesson was rather tight and some of them admit that they once or twice exceeded this time-limit.

Teacher TS1 is a person who likes to discuss a theme in depth, especially if it is a topic he is interested in, like the ETA (lesson 4). Normally, he would at least have been talking about this topic for half an hour. During the lessons of the experiment, he felt obliged to limit himself to the five minutes scheduled for this pre-reading activity. Normally he did not start a new subject, unless he was sure that his students had acquired the necessary knowledge and skills to deal with it. Forty-five minutes were too short to be able to do this during the lessons of the training programme. He especially regretted the fact that there had hardly been time for feedback on the tasks of the training programme, the most important aspect of the programme. Because such a situation did not combine with his character and his normal attitude towards the students, he had sometimes decided to dedicate more time to the tasks than was scheduled.

Also for Teacher TS2, 45 minutes were not sufficient to really profit from the lessons of the training programme. Mostly, his students had not even understand the main points of the text after these 45 minutes. Therefore, the students could not internalize the words of the texts. Especially the weaker students had suffered from this situation. Despite all this, he had never decided to prolong the lessons of the training programme.

The teachers who executed the English training programme found that the time-table had not always been realistic. They sometimes had to hurry to finish the lessons of the programmes in time. Mostly they succeeded, but Teacher TE3 once or twice had to go on for five or ten more minutes. Teacher TE1 mentioned an original solution. She had used an alarm clock in the classroom to be sure that she would not overstep the time-limits.

Concerning the teachers of the control groups, Teacher CS1+CS2 had sometimes needed five or ten extra minutes, whereas Teacher CE1+CE2 had always managed to perform the lessons within the 45 minutes.

Time-tables

Teacher TS2 explained that, at the time of the experiment, his schedule had not included a two-hour block for his Spanish course. The lessons of the programme had to take place in the first, seventh or eighth period. This was an unfavourable situation because students either arrived late or went away earlier to catch their bus. It would have been preferable to have had a two-hour block (90 minutes), like the other six teachers. Teacher TS2 explained that these circumstances had made it difficult to plan the context and reading comprehension tests of the experiment, for which about 60 minutes were planned.

Homework

We had planned that the programmes would not contain homework assignments. Nevertheless, the interviews revealed that some of the teachers had not kept to this agreement.

The students of Teacher TS1 had studied the schemes (overviews of context rules and logical connectors; cf. 4.6) at home. The students of Teacher TE1 had to learn the words and logical connectors (cf. 4.6). Finally, the students of the two school classes that together formed the control group for Spanish had to study the vocabulary of the texts of the control programme.

Planning

The teachers of the Spanish training groups (Teacher TS1 and Teacher TS2) were not satisfied with the planning of the experiment as a whole. They strongly believed that it would have been better to spread the programme over a larger period, e.g. once a month instead of once a week. During the four months of the experiment too much time was dedicated to the teaching of reading.

The other teachers, however, were positive about the intensive seven-week programme (Teacher TE2, Teacher CS1+CS2 and Teacher CE1+CE2) or did not have a preference.

Foreign language proficiency level

According to the teachers, the foreign language proficiency level of the students had sometimes affected the execution of the training and control programmes.

The teachers of the Spanish training groups often referred to the low foreign language proficiency level of their students. They partly traced the difficulties they had been confronted with during the execution of the training programme to the new textbook used for the Spanish course. This textbook was based on a functional-notional approach towards foreign language teaching. In the lessons preceding the experiment, the students had not practised many reading tasks, but mainly oral dialogues in Spanish. During the execution of our experiment this often led to comprehension problems.

Teacher CS1+CS2 reported that the students of CS1 had coped better with the programme than those of CS2, which was a rather weak groups of students.

Motivation

The programmes had a discouraging effect on the students of the two Spanish training classes. Teacher TS1 had never sufficient time to really explain to his students what the training programme was about. In the course of the training programme, his students began to dislike the lessons because they did not feel that they benefited from the training programme. They knew how to perform the task, had learnt the 'context rules' by heart, but were not able to understand the local contextual information. They were used to knowing exactly what was expected from them, what they had to prepare at home and what they had to study to pass the examinations. The lessons of the programme were not like this at all. This made them feel insecure.

Also the students of Teacher TS2 lost the motivation to perform the lessons of the training programme. According to their teacher, TS2 was a very weak group. The students obviously had a hard time dealing with the exercises in the use of context. A small group of students did not agree at all with the contents of the training programme. They found it nonsense, useless and very vague. This made that it was difficult to execute the training programme.

Finally, the teachers of the English training programme reported that the students had been somewhat less motivated at the end of the programme. This was probably due to the fact that they got used to the tasks of the training programme.

The control programmes

Discussing the pilot in Chapter 4, we saw that the control condition sometimes had included aspects of context use. Some of the teachers who had executed the control reading programmes during the pilot briefly introduced text-topic or made students aware of the loan words and cognates of a reading text. We had taken measures to avoid that the lessons of the control condition of the main experiment would include such aspects of the use of context: the tasks of the control reading programmes did not embrace pre-reading activities and tasks in which students had to deal with loan words and cognates (cf. 4.8).

Despite the measures, Teacher CS1+CS2 often referred to loan words and cognates in the lessons of the control condition for Spanish. Teacher CS1+CS2 mentioned that languages and their interrelationships was in fact his hobby. Therefore, he often used other languages to explain Spanish words. He believed that he did it too frequently and that his students were sometimes bored with this. He also 'saw' relationships between words that were rather far-fetched. For example, dealing with the Spanish word 'someter' (= to subdue, to restrain, to suppress) he told his students that, to remember the meaning of this word, they could think of the English verb to 'smite' (= to hit hard), which in its turn, sounds like the German word 'schmeissen' (= to throw). He even mentioned an example of a 'related' Russian word.

Also the lessons of Teacher CE1+CE2 sometimes contained activities that can be classified as context use. This teacher often drew the students' attention to the word class of an unknown word. Teacher CE1+CE2 gave an example. In one of the lessons of commercial correspondence the following sentence appeared: "They

invoice him monthly". The meaning of the word 'invoice' was not familiar to the students. He described the meaning of this word in English and one of the students came upon the meaning 'bill'. Teacher CE1+CE2 explained that 'invoice' had to be a verb, because it was preceded by 'they'. This made it easier for the students to find the right solution.

III. EFFECTS

Despite the problems some of them had been confronted with, each of the five teachers of the training programmes mentioned some positive effects of the training programme in the use of context. The main effect of the training programme was that the students had learned that they did not necessarily have to know each single word of a reading text. They had become more independent and were not afraid to be confronted with unknown words in a reading text. Moreover, they were aware of the function of the logical connectors and used this knowledge while reading. Most teachers still applied aspects of the training programme in the use of context in their teaching after the experiment. Teacher TS1 continued dedicating attention to context-use strategies. Teacher TE1 let the students perform gap-filling exercises, made by herself. Both Teacher TE2 and Teacher TE3 often referred to the training programme, especially to what had been learned about logical connectors. Only Teacher TS2 was not sure whether he would use aspects of the training programme in his further teaching.

6.2.2. Instruments

In this section the teachers' evaluations of tests (I) and questionnaires (II) are presented.

I. TESTS

Both for the pre- and post-test sessions, three different types of tests were used to measure the students' foreign language skills, i.e. vocabulary tests, context tests and reading comprehension tests (cf. 4.8). The teachers' comments are subdivided into contents, execution and test results.

Contents

Both teachers of the Spanish training group (Teacher TS1 and Teacher TS2) found that the level of the tests had been rather high regarding the foreign language proficiency of the students. The teachers who had executed the English training programme in the use of context found the tests suitable. Finally, the two teachers of the control condition found the tests rather difficult. This was partly due to the fact that, contrary to the students who had followed a training programme in the use of context, their students were not familiar with the task of the context test, i.e. *Guess the meaning of the unknown word* and, therefore, sometimes felt insecure about how to deal with this type of testing.

Execution

The teachers had the impression that their students had performed the tests of the experiment seriously. Teacher TS2, however, was not completely satisfied with the execution of the last test of the experiment. The students had to make this test - the reading comprehension post-test - under unfavourable conditions. The time-table did not provide for a two hour block of Spanish (cf. 6.3.1, II. Execution, Time-table). For the context and reading comprehension tests, however, more time was needed than 45 minutes. Therefore, during the test periods Teacher TS2 had usually asked a colleague to give the students the opportunity to complete the tests in the following lesson of the students' time-table. For this last test, this had not been possible. The only solution was to perform the test in the last hour of the time-table and give the students the opportunity to finish the test afterwards (late afternoon). This meant that the students had to perform the test partly in their spare time. They did not really enjoy this and tried to finish as soon as possible.

Test results

At the beginning of the experiment, we agreed with the teachers that the test results would partly determine the official marks for Spanish and English of that particular study year. However, after the execution of the experiment, only three of them decided to actually do so. One of these teachers had performed a training programme in the use of context (Teacher TS2), the other two had executed the control programmes (Teacher CS1+CS2 and Teacher CE1+CE2). The latter reported that the results were quite as expected. The students that normally performed well, also performed well in the tests of the experiment.

For Teacher TS1 the reason *not* to include the tests results in the final marks was his dissatisfaction with the execution of the experiment. His students felt like they lost control over their foreign language performance, that they had not improved their Spanish at all and had not known how to prepare themselves for the tests. Therefore, he found that it would not be appropriate to include the test results in the marks of the Spanish course. He had included, however, his personal evaluation of the degree of the students' active participation during the lessons of the training programme.

The teachers of the English training programme gave two reasons for *not* including the test results in the official marks. Firstly, they felt that they could not blame the students for performing poorly on the tests of the experiment. Secondly, the fact that some students had not made all tests of our experiment, made it sometimes difficult to compare the overall results.

II. QUESTIONNAIRES

During the post-tests, the students filled in four questionnaires (cf. 4.9.4). Below, the teachers' observations about contents and execution are reviewed.

Contents

On the whole, the teachers had not very many comments with regard to the contents of the questionnaires.

Two teachers, i.e. Teacher TS2 and Teacher TE2, observed that the students did not like the control questions. The students considered it nonsense to have to answer the same kind of question twice and interpreted this as not being taken seriously.

Teacher TE3 refers to a different kind of problem related to the contents of the questionnaires. She remarked that the students did not feel comfortable with the more personal questions. Examples are the questions of the questionnaire *Self-confidence* and questions like "How well do you think you performed this reading task" (questionnaire *Reading Ability*). She explained that students found it always hard to judge themselves.

Execution

Most teachers did not mention problems with regard to the execution of the questionnaires.

However, Teacher TS2 and Teacher TE2 mentioned that some students had not liked to fill in the questionnaires and always complained about it.

Additionally, Teacher TE1 observed that her students sometimes had made jokes about the contents of the questionnaires. For example, they had remarked that the negative scale was not large enough. Nevertheless, she was under the impression that, once they were filling in the questions, they took it seriously. She had noticed that some of her students also wrote down personal remarks and comments. This was not an obligatory part of the questionnaires.

6.3. Summary and conclusion

The results of the *students' evaluations* showed that the students were not convincingly positive about the training programmes in the use of context.

The Spanish training group doubted whether the training programme had positively influenced their ability to contextually guess the meaning of unknown words. They had not enjoyed working with the training programme, concluded that there had not been much variety in the tasks, were not really positive about the teaching during the experiment and did not believe that the training had helped them to better understand Spanish texts. The analysis by school class showed that TS1 tended to be less negative about some of these aspects than TS2.

The English training group was somewhat more positive about the training programme. The majority of the students had enjoyed the lessons of the programme. The students of two school classes, i.e. TE2 and TE3, found that the training programme had been a helpful tool to learn how to guess the meaning of unknown words, that the aim of the programmes had been clear, that the results of the tasks had been satisfactory and that there had been enough time to read the texts and to perform the tasks. Most of the students of TE3 were also positive about aspects such as pleasure experienced during the execution of the training, the variety in the tasks of the programme, the teaching during the experiment and the improvement of their reading comprehension ability. Different results were found for the third school class that had participated in the English training, i.e. TE1. This class tended to be

negative about almost all aspects of the training programme in the use of context.

The students were relatively positive about the tests of the experiment. On the whole, the aim of the tests had been clear, there had been enough time to make the tests and the results were more or less satisfactory.

The *teachers' evaluations* of the programmes and tests provided explanations for some of the outcomes of our experiment. In the interviews, some possible reasons of the low proficiency level of the training groups for Spanish were given. In Chapter 5 we saw that, already before the experiment started, there were substantial differences between the Spanish training and control groups. According to the teachers, it was not unthinkable that the recently introduced textbook, based on a communicative approach to foreign language learning, had not been the suitable preparation for an intensive reading programme. This new textbook had been lexically and grammatically analysed to be sure that the Spanish training programme linked up with the students' proficiency level. This, however, could not compensate for the fact that the students had been less experienced in reading and analysing texts than the students of the control condition, who used a rather traditional text book, based on the grammar-translation method. The lack of reading skills negatively influenced the execution of the training programme. This was confirmed by the negative evaluation of the programme by the students and the remarks made by the teachers in the interviews. The fact that the tasks and tests were often (too) difficult for the students might have led to a decreasing motivation.

This probably also affected the two teachers of the Spanish training group, i.e. Teacher TS1 and Teacher TS2. They strongly felt how their own teaching objectives clashed with the objectives of the experiment. More than once, these teachers explained that they felt guilty towards their students because they could not give them the support they needed. And, although they had agreed to follow our instructions and tried not to show these negative feelings in their lessons, they probably did not totally succeed in hiding their irritations.

Apparently, TS2 suffered more from this situation than TS1. The students evaluated the training programme very negatively and the teacher referred to the fact that this particular school class had always been very weak. Interpreting the results of the interviews, however, one has to consider that, at the time of the interviews, all teachers had already been informed about the results of the post-tests. The fact that the teachers of the Spanish training programme knew that their students had not performed well, might have influenced their opinions about the effectiveness of the programmes and tests.

There are reasons to believe that the students of TS2 had not seriously performed the reading comprehension post-test, because this test had to be partly executed in their spare time.

It was inevitable that the students had been aware of the fact that they were participating in an experiment: they had to perform more tests than they were used to, they normally did not perform multiple-choice tests and during the experiment they did not work with their regular foreign language textbooks.

Our efforts to control certain variables had not always been successful. Some of

the teachers exceeded the time planned to perform a lesson, set homework assignments or decided, afterwards, *not* to incorporate the results of the tests of the experiments in the final marks.

The teachers involved in the training programme in the use of context were convinced of the potential of the training programme. They estimated that their students had become more confident and independent readers, knew the meaning and function of logical connectors in a text and had learned some useful strategies to infer the meaning of unknown words. Both from the students' and teachers' perspectives, the appreciation of the English training programme tended to be more positive than of the Spanish training programme.

In Chapter 4, a deliberate choice was made to carry out a field-experiment. The main reason for this decision was the conviction that teaching programmes such as the training programmes under study, had to be tested in an ecologically valid context, i.e. the foreign language classroom itself. This meant that all kinds of measures had to be taken to control other variables that may play a role in such a situation. These measures are reflected in the choice of school-type and subjects, the efforts to link up with the foreign language curriculum and the students' level of foreign language proficiency, the frequent contacts with the teachers who executed the programmes, the strict instructions, the selection of reliable instruments and the execution of a pilot experiment.

The results in this chapter, however, reveal that despite these measures, there are some factors that may have affected the reliability of some of our research results. In the last chapter of this book these factors will be explored and conclusions will be drawn.

Chapter 7

Discussion

In the previous chapters the quantitative and qualitative research results were presented. In this last chapter conclusions are drawn from our study of the effects of the training programmes in the use of context for beginning learners of Spanish and intermediate learners of English. Section 7.1 is a summary of the research results outlined in Chapters 5 and 6. Section 7.2 is the discussion of these results. This section aims at suggesting explanations for the outcomes of the experiment. The implications of our findings for the teaching of foreign language reading comprehension are discussed in 7.3. In the last section of this chapter suggestions for further research are given.

7.1. Summary of results

In Chapter 4 of this study three research questions were formulated. The first question was whether a training programme in the use of context would be more successful than a more traditional way of dealing with texts in the classroom, both with respect to the independent variable context use and the dependent variable reading comprehension. The second was whether there would be differential results of such a training programme at an elementary level of foreign language proficiency (Spanish) and an intermediate level (English). The third research question was whether four reader-dependent variables, i.e. vocabulary knowledge, attitude towards the foreign language, self-confidence and the student's estimate of the degree to which he or she actually uses context to infer the meaning of unknown words, would be relevant for context use and for reading comprehension.

Research questions 1 & 2: Effects of training the use of context

Chapter 5 dealt with the effects of training context use. The overall quantitative results of the experiment did not convincingly prove that the two training programmes in the use of context had been successful. Both for Spanish and English, the subjects that followed the training programmes in the use of context had a significantly higher mean gain score on the context tests than those who followed a control programme. Nevertheless, partially contrary results were established after comparing the mean gain scores on the reading comprehension tests. In the case of Spanish, the subjects of the training condition performed significantly poorer than the control group and for English no significant differences

between training and control conditions were found.

Therefore, the answer to the first research question was negative. The results of our project did not prove that our training programmes in the use of context had been more successful than the control programmes. Although the mean gain scores of the training groups for Spanish and English on the variable context use appeared to be significantly higher than the mean gain scores of the control groups, there were no significantly positive effects on the variable reading comprehension.

In the light of these results, we could not answer the second research question. It was not possible to establish whether the training programme in the use of context for beginning learners of Spanish had been more successful than the training programme in the use of context for intermediate learners of English or vice versa.

Although a comparison of overall mean scores is in line with the research design presented in Chapter 4, there were serious reasons to believe that the nature of the data was too complex to be properly described by this overall quantitative analysis. The results of the pre-tests showed that there were significant differences between the Spanish training and control groups, already before the experiment started. Moreover, the control group's mean score on the context test of the pre-test was very high, which could be an indication of a ceiling effect. Finally, except for the performance on context use of the Spanish training group, all mean gain scores were negative.

Therefore, two additional quantitative analyses were presented in Chapter 5: an analysis by school class and one at the individual subjects' level.

The analysis by school class aimed at establishing whether subgroups, i.e. school classes, performed differently from the overall performance of the group. The results showed that only part of the school classes showed significant positive effects of the training programme in the use of context.

The aim of the analysis of the performance of each subject was to establish whether subjects behaved differently from previously established reference-values, representing the performance of the subjects of the control condition. The results of this analysis showed that, both for context use and reading comprehension, extreme deviations from the reference-values were almost exclusively found with the subjects that had followed the Spanish or English training programme in the use of context.

An important outcome of the two additional analyses was that there were differences between the performance of school classes and of students. Concerning the latter, no further research results were available because no qualitative data had been gathered of how individual students dealt with context-use training. In several occasions, however, differences between school classes were analysed.

Below, the results of our study are summarized for each school class that followed a training programme in the use of context. This summary embraces the quantitative

results by school class, outlined in Chapter 5 and some of the results of the students' and teachers' evaluation of the project, such as outlined in Chapter 6. A distinction is made between Spanish and English.

Spanish

To establish possible effects of the training programme in the use of context for beginning students of Spanish, the results of the two classes that followed the training programme (TS1 and TS2) are compared with the results of the two school classes of the control condition (CS1 and CS2).

Spanish training programme - school class 1 (TS1)

The results of the quantitative analysis by class, showed that TS1 scored very low in pre-tests sessions, whereas the control classes (especially CS1) scored very high. TS1 is the only school class for Spanish that scored better in the context post-test than in the pre-test, leading to a significant difference between the mean gain scores of TS1 and those of the two control classes. However, in view of the differences in the pre-test scores, this positive finding could not indisputably confirm that the Spanish training programme in the use of context had been successful. Moreover, the mean gain score of TS1 on reading comprehension was not significantly different from the results of CS1 and CS2.

The students of this class were not positive about the training programme. According to them, the programme had neither been useful to learn how to contextually infer the meaning of unknown words, nor to better understand Spanish texts.

The teacher of this school class was both positive and negative about the effects of the training programme. On the one hand, he observed that his students had learned how to work independently and had become more autonomous readers. On the other hand, he concluded that the students' proficiency level of Spanish had been too low to actually benefit from the strategies learned during the training programme. According to the teacher, this had led to frustrations and decreasing motivation.

Spanish training programme - school class 2 (TS2)

No significant differences were found between the mean gain scores on context use of class TS2 and control classes CS1 and CS2. Considering performance on reading comprehension, the results revealed that TS2 scored extremely low on the last reading comprehension test of the experiment. This led to a mean gain score of -.36. This mean gain score was significantly poorer ($p < .01$) than the mean gain scores of the two control classes.

The students of TS2 evaluated the programme and tests even more negatively than the students of TS1. They found the programme ineffective, unpleasant to work with and were dissatisfied with the teaching during the experiment. Often, they had not had enough time to read the texts and perform the tests.

According to the teacher of this school class, this specific group was very weak and had experienced many problems during the execution of the experiment. Like

the students of TS1, this class was not motivated to execute the training programme and tests. Additionally, the time-table of the Spanish course was unfavourable. This is the reason why the students of TS2 had to make part of the reading comprehension post-test in their spare time. Because they did not like this situation, their teacher suspected that they had not seriously performed this last test of the experiment. Nevertheless, he noticed that his students had acquired a different attitude towards reading tasks. They were more willing to try to solve comprehension problems by themselves.

English

Three school classes executed the training programme in the use of context for intermediate students of English: TE1, TE2 and TE3. This last class is not dealt with in this summary, because after eliminating the missing data, the group consisted of only four subjects. The results in the previous chapter, however, revealed that the evaluation of the project by the students of this school class was positive and that also their teacher positively evaluated the lessons of the training programme in the use of context. CE1 and CE2 were the two control classes for English.

English training programme - school class 1 (TE1)

The quantitative results by school class showed a significantly better mean gain score of TE1 on context use, in comparison with the mean gains of CE1 and CE2 (TE1 vs. CE1: $p \leq .05$; TE1 vs. CE2: $p \leq .01$). For reading, no significant differences were found between TE1 and the control classes.

The students' evaluation of the project was predominantly negative. They believed that they had not learnt much from the training programme in the use of context, they had not liked the lessons of the programme and had not been satisfied with the teaching during the experiment. T-tests confirmed that the total score of this class on the questionnaire *Evaluation of Programmes and Tests* was significantly lower than the score of TE2 and TE3, the two other classes that had followed our English training programme ($p \leq 0.01$).

The teacher of this group, however, was satisfied with the effects of the training programme. She referred to the fact that students had become more independent and did not easily panic if a text contained unknown words.

English training programme - school class 2 (TE2)

There were no significant differences between the mean gain scores of class TE2 and the outcomes of the two control classes.

The evaluation of the students of TE2 was mainly positive. They found that the training programme had taught them how to contextually guess words, although they were not sure whether it had improved their reading comprehension ability as well.

The teacher of this school class was positive about the effects of training the use of context. He believed that his students had learned how to deal with logical connectors and context strategies and confirmed that they had become more self-reliant while reading.

Conclusions

This review shows that, although there was hardly any quantitative evidence of positive effects of the training programmes in the use of context, there was one important overall qualitative result. Each of the five teachers that had executed one of the training programmes in the use of context recognized that his or her students had become more autonomous. They were less dependent on their teacher's help and were less afraid of being faced with comprehension problems. This, however, could not be confirmed by other data: we did not measure the students' autonomy before and after the experiment.

On the whole, the execution of the English training programme had been less problematic than the execution of the Spanish training programme. The teachers were positive about the project and had not faced serious problems during the lessons of the training programme in the use of context. There were two positive outcomes of the English training programme. Firstly, TE1 had a significantly higher mean gain score on context use than the control classes. Secondly, according to the judgement of the students of TE2 and TE3, the training programme had been useful to learn how to contextually guess words.

The major part of the problems of the experiment concerned the execution of the Spanish training programme. According to the two teachers involved, these problems were caused by the low Spanish proficiency level of their students. The students were not able to successfully perform the tasks of the training programme, did not experience any progress and lost their motivation to participate. The results of the students' evaluation and the quantitative results by class revealed that TS1 had dealt with this situation better than TS2. TS1 evaluated the training programme less negatively than the students of TS2 and showed positive mean gains on context use and reading comprehension. TS2 did not show any positive effect of the training programme at all and even scored significantly lower on reading comprehension than the control classes.

The analysis by school class did not only reveal differences between classes that had been exposed to one of the two training programmes in the use of context, but also between classes of the control condition. For Spanish, CS1 had a higher mean gain score on reading comprehension than CS2. Moreover, the students' evaluation showed that the total score of the students of CS1 was significantly higher than of CS2 ($p < 0.01$). The teacher of these two classes confirmed that the students of CS1 were better than the students of CS2. An interesting outcome of the students' evaluation was that the major part of the students of CS1 had found their control programme useful to learn how to contextually guess the meaning of unknown words.

These outcomes are different from what was found in the 1987-project. The results of this follow-up study do not confirm that the training programme in the use of

context for beginning learners of Spanish significantly improved context use and reading comprehension. On the contrary, in this case, context-use training at a low level of foreign language proficiency led to negative effects, such as frustration and lack of motivation. For English, no such comparison can be made for the simple fact that the 1987-project only involved Spanish as a foreign language.

Research question 3: Relevance of four reader-dependent variables

The third research question dealt with the relationship between four reader-dependent variables on the one hand and context use and reading comprehension on the other. These variables are vocabulary knowledge, attitude towards the foreign language, self-confidence and the student's estimate of the degree to which he or she actually uses context. To measure these four reader-dependent variables the post-test included a vocabulary test and a set of questionnaires.

The results in Chapter 5 showed that there were significant correlations between the results of the vocabulary test and the context test on the one hand, and between vocabulary knowledge and reading comprehension on the other. For Spanish, the correlation between vocabulary and reading comprehension was low ($.35, p \leq .05$).

Conflicting results were found for the variable *attitude*. Concerning Spanish, there were significant correlations between this factor and performance on context use and reading comprehension. In the case of English, however, this factor did not significantly correlate with the students' performance on the context and reading comprehension tests.

No significant correlations were found between the variable *estimate of context use* and the performance on the context or reading comprehension tests.

Finally, we found no significant correlations between the students' reported level of *self-confidence* and their actual test results, except for one. The variable 'self-confidence' significantly correlated with the results of the Spanish context test. This correlation, however, was rather low ($r = .28; p \leq .05$).

Conclusions

'Foreign language vocabulary knowledge' was a relevant variable in this study. The variable 'attitude' only was important in the case of beginning learners of Spanish. The variable 'estimate of context use' did not play a significant part. Finally, 'self-confidence' only was relevant in the case of the Spanish context test.

Considering Spanish, both in our study and in the previous project, the vocabulary test showed significant correlations with the context and reading comprehension tests. For context use, the correlation coefficient found in this study was comparable with what was found in the 1987-project. For reading, the correlation found in our project was lower than in the 1987-project. Strictly speaking, however, these results cannot be compared: in the current project vocabulary was measured after the experimental treatments and in the previous project before the execution of the training and control programmes.

Neither in the case of the three variables *attitude*, *estimate of context use* and *self-confidence* it is legitimate to compare the results of the current study with the results of the earlier project. As outlined in Chapter 5, our follow-up was organized somewhat differently than the 1987-project. In the latter project, the subjects filled in questionnaires both before and after the programme, whereas in our experiment only in one of the post-test sessions. Therefore, it was not possible to establish differences between performance on the questionnaires before and after the programmes, such as had been the case in the 1987-project. Finally, in the 1987-project a different statistical technique was used. However, we may conclude that by means of both our analysis and the analysis carried out in the 1987-project, 'attitude' appears to be relevant for beginning readers of Spanish and the students' estimate of context use, reported in the questionnaire 'Evaluation', inconsistent with their actual performance on the context tests.

7.2. Discussion of results

In Chapter 4 we explained that ecological validity was one of our main concerns. The central question of this study is whether context-use training is useful to develop foreign language reading skills. We believed that the best way to explore the effects of a training programme in the use of context would be to use it in the conditions for which it is meant, i.e. a classroom setting. Therefore, we conducted a field-research and not a fully controlled laboratory-like experiment. A field experiment is always risky. In our case, the experiment was a follow-up and amplification of an earlier, small-scale field-experiment on the effects of training the use of context, involving only one foreign language (Spanish), one school, one teacher, a limited number of subjects and training and control conditions which only contrasted with regard to one aspect (teacher-guided versus learning individually, i.e. without supporting exercises, feedback or any further explanation of the teacher). The subjects of this earlier project were allotted to training and control conditions on the basis of the results on a context test. The weak context users followed the teacher-guided training programme, whereas the strong context users followed the control programme. Because we wanted to assess whether we could also find positive effects of training context use on a larger scale and under less strict conditions, we decided to involve a second language (English) and create a more obvious difference between training and control conditions. The results of our experiment, outlined in the previous sections, show the complex interaction of variables in a classroom setting, which makes it often difficult to avoid contamination. In this section we will suggest explanations for the outcomes of this study. In the first sub-section the results on the context and reading comprehension tests are discussed (research questions 1 & 2), in the second part the results of the analyses of correlations between four reader-dependent variables on the one hand, and context use and reading comprehension on the other (research question 3).

7.2.1. Training context use

Our research results could not provide a clear answer to the question whether training context use is relevant for reading a foreign language. There were several factors that may have contaminated the results of the experiment. These factors will be discussed below.

Reliability of the results of the Spanish training and control groups

There were three reasons to question the reliability of the results of the Spanish groups: regarding the control group, we suspected a ceiling effect, the training group experienced a decreasing motivation and, finally, there were reasons to believe that one of the school classes of the training group failed to seriously perform the reading comprehension post-test.

Ceiling effect

The pre-test results, revealed in Chapter 5, showed that the training and control groups for Spanish were not as comparable as assumed. In the context pre-test the control group for Spanish scored more than twice as high than the training group and a ceiling effect could not be excluded. In the case of reading, the difference was less substantial. This can probably be explained by the types of tests. The reading comprehension test is a multiple-choice test and the context test an open test. It is likely that a multiple-choice test is less sensitive than an open test (see also, Choice of instruments, p.115). A competing explanation could be that the students of the control condition had cribbed while making the context pre-test. The latter, however, is not very likely, because the Spanish control group also scored relatively high in the context post-test, in comparison with the Spanish training group (cf. Chapter 5, Table 5.1).

Motivation

There were strong reasons to believe that the results of the two school classes of the experimental condition were negatively influenced by a decreasing motivation. Both our classroom observations and the results of the teachers' interviews showed that, in the course of the execution of the experiment, the Spanish students got less and less motivated to cooperate in the lessons of the training programme because they did not experience any progress. According to the teachers this was due to their low proficiency level of Spanish.

Unreliable test results

The teachers' evaluation in Chapter 6 showed that the teacher concerned suspected that TS2 did not seriously perform the reading comprehension post-test. To check whether this in fact had been the case, we conducted an analysis of the reliability of the results of TS2 on this particular test. This analysis showed a negative reliability coefficient of $-.06$. Consequently, these data are not reliable. The reason for this unreliable performance is probably a combination of lack of motivation and the unfavourable time-table of this school class, which made that the students had to make the test partly in their spare time.

We also checked whether the outcomes of other tests and/or other school classes during the execution of the experiment were influenced by contaminating factors. Besides the results on context use of control class CS1 (ceiling effect) there were no reasons to suspect this.

Considering all this, a comparison between the Spanish training and control conditions will not be advisable. A comparison of mean gain scores on the context tests is probably contaminated by a ceiling effect, the low proficiency level of the training group for Spanish negatively influenced the execution of the training programme and TS2 did not seriously perform the reading comprehension post-test. These contaminating factors also explain the apparent contradiction between the overall results of the Spanish training group.

In Chapters 4 and 5 we already anticipated that a comparison of the results of the Spanish training and control groups would not be legitimate. In Chapter 4, we explained why we decided not to stop the execution of the experiment, although we knew the results of the pre-test sessions. Our arguments were mainly practical. It would be hard to find a 'fresh' suitable population for our experiment, the school in Kleve had also been involved in the 1987-project, a pilot experiment had already been executed, the programmes and tests had been adjusted to the foreign language curriculum and textbooks of the two schools, the school-boards had given their permission for the execution of the experiment and there was no reason to believe that for English we would be confronted with the same type of problems.

Selection of subjects

The performance of the Spanish training and control groups on the pre-tests can partly be explained by the selection of subjects. The main criterion to select the subjects that participated in our project was that they were second-year students of a certain type of vocational training in Germany. Two schools were involved in our experiment. One of these schools was the school where also the 1987-project had taken place. In our follow-up experiment, the subjects of the last-mentioned school were allotted to the training condition. The lessons of the control programmes were executed at the second school. We assumed that the level of foreign language proficiency of the students of the two schools would be comparable, considering the fact that the students were involved in the same type of vocational education. Moreover, the results of the pilot experiment had not given reasons to suspect that this might not be the case (cf. 4.5).

In the teachers' interview, a possible explanation for the apparent difference between the level of Spanish at the two schools of our experiment was given. One year before the experiment started, for the first time communicative teaching materials were used at the school where the training programmes were executed. At the second school a more traditional textbook was used. The 'new' communicative text book aimed at developing verbal communication skills, not at training reading comprehension skills. According to the teachers of the training classes, their students had been less exposed to reading and analysing texts than the students of

the control condition and were, therefore, less skilled in performing reading comprehension tasks. The quantitative results by school class, however, show that this explanation is not likely to be valid (cf. Appendix XIV, Tables 1 & 3). Although both control classes CS1 and CS2 scored better than the two training classes on the context pre-test, this is not the case for reading. On the reading comprehension pre-test, the score of TS2 is comparable with the scores of both CS1 and CS2. This indicates that the communicative textbook cannot be the only reason that the students of the training condition performed poorly on reading comprehension. It may well be that not only did the variable 'textbook' play a part in our research, but probably also other characteristics of school classes, students and teachers.

On the basis of the data gathered in this study, it is not possible to determine with any certainty what factors may have contributed to the different starting-situations of the Spanish training and control groups.

Differences between school classes

The four groups of the research design were not randomly selected but consisted of already established school classes at the two schools involved in the experiment (cf. 4.4). The pre-test results by school class, outlined in Appendix XIV demonstrate that, both in the experimental and control conditions, there were sometimes differences between classes of the same condition, at the beginning of the experiment. TS1 tended to score lower on the pre-tests than TS2, TE1 tended to perform poorer than TE2 and CS1 tended to score higher than CS2. In the case of the reading comprehension pre-tests these results appeared to be significantly different (Spanish: $p \leq .01$; English: $p \leq .01$). Additionally, the scores on context use of the two Spanish control classes were significantly different ($p \leq .01$).

These differences between school classes were partly confirmed by the results of the teachers' interviews in Chapter 6. Teacher CS1+CS2, for example, revealed that CS2 was a weaker group than CS1. For Spanish, however, the results of the teachers' interviews sometimes appear to be rather subjective. According to Teacher TS2, this school class was a very weak group of students. This was in contradiction with what was found in Chapter 5. As outlined above, TS2 performed better on the pre-tests than TS1 and, with respect to the reading comprehension pre-test, equally well as the two control classes.

The results in Chapter 5 revealed that the variable school class was very important in this research. For English, we only found a significant effect of our training programme in the use of context for one school class. An interesting finding is that we also found a significant difference between the mean gain scores of the

in the use of context. In this study, however, we cannot give a satisfactory explanation for individual differences in the effects of training the use of context because we have not gathered data on how individual students worked with our programmes.

Differences between teachers

The ecological validity of the investigation required that the training and control programmes would be executed by the same teachers that would normally be responsible for the Spanish and English lessons of the subjects. Despite our attempts to control variation between teachers, the role of individual teachers was sometimes manifest. Moreover, our deliberate choice for a control condition in which the teachers were asked to deal with texts as they would normally do, implied that we could not fully control how the lessons of the control condition would be executed. To control the situation as much as possible, we decided to involve only one teacher for Spanish and one for English.

A total of seven teachers collaborated in the research project. The training programme for Spanish was carried out by two teachers, the training programme for English by three. In other words, each single class of the training condition had its own teacher. In the control condition only two teachers executed the programmes. One teacher was involved in the lessons of the two classes for Spanish and the other one in the English lessons of the two control classes for English. Below we will discuss possible consequences of the teacher variable. We will refer to our teacher instructions, the commitment of the teachers and the extent to which we succeeded in managing the execution of training and control programmes.

We tried to control differences between teachers as much as possible by giving them a clear explanation of the aim of the training or control programmes and written instructions for each of the seven lessons. These written instructions included a strict planning of and also possible solutions to the exercises. The results of this study cannot give a definite answer to the question whether these instructions had been sufficient.

There was no time left to train the teachers before the experiment started. The school-boards had only given us permission to intervene during the first semester and we needed this period to execute the pre-tests, programmes and post-tests. Three of these teachers, however, were not unfamiliar with the programmes because they had also participated in the pilot, which took place one year before the experiment. The teachers that had not participated in this pilot were the three teachers of the English training programme in the use of context and one of the teachers of the Spanish training programme, i.e. Teacher TS2. There is, therefore, variation between the teachers in the degree of previous experience with one of our training or control programmes. In the case of English, the three teachers of the training condition had not been involved in the pilot, whereas the teacher of the control condition had also participated in the tryout. For Spanish, Teacher TS1 was a more

experienced context-use trainer than Teacher TS2. The first-mentioned teacher not only had been involved in our pilot study, but also in the pilot and experiment of the 1987-project. The fact that the mean gain scores of TS1 were positive, both in the case of context use and reading comprehension, might indicate that, despite the problems during the execution of the experiment, this school class did benefit from the training programme (cf. Appendix XIV).

It would have been possible to avoid the variation in previous experience, but only at the expense of the number of subjects involved. Because after the pilot, the fourth research question - establishing whether there would be a transfer effect of a training programme in context-use in Spanish to the performance on English, or vice versa (cf. 4.3) - was dropped, it was no longer a prerequisite that the subjects would be studying both Spanish and English. This meant that a larger group of students could cooperate in the experiment. Because we found it important to have a large number of subjects, we decided to select also the students of teachers without previous experience with our programmes.

The teachers involved in our experiment had deliberately chosen to participate in the experiment. The five teachers involved in the training believed that the programme was refreshing and declared that they were convinced of its practical use. For the three teachers involved in the English training programme, this positive attitude did not change in the course of the programme. In the case of Spanish this was different. Once confronted with the fact that the execution of the training programme caused some problems, the two teachers adopted a less positive attitude towards the project. In the interviews they both declared that they were very glad that the experiment had come to an end and that they had returned to their everyday classroom routine. From the results in Chapter 6, we concluded that TS2 suffered more from this situation than TS1. This can probably be explained by the fact that Teacher TS2 was less experienced in the execution of context-use training than Teacher TS1. The conclusion we can draw is that, in our experiment, commitment of teachers was strongly related to the success of the intervention in their classroom.

Contents of training and control programmes

Earlier in this study we paid ample attention to the differences between the 1987-project and this follow-up (cf. 4.2). In the 1987-project the tasks of the control treatment were not really different from the tasks of the experimental treatment. Then, the surplus-value of the experimental condition consisted in the systematic training in the use of context through supporting exercises, schemes and feedback of the teacher, whereas the subjects allotted to the control condition had to perform these tasks individually and without any help. We hypothesized that it was not unthinkable that the significant positive influence of the context-use training programme, found in the 1987-experiment, reflected (also) the positive effects of the availability of explicit help from a teacher and not (only) the differences in context use and text comprehension. Therefore, we intended to manipulate a sharper contrast between training and control conditions. In this follow-up we assessed the

effects of two types of reading training, i.e. training context use (cf. 4.7.3) and a control programme that reflected how teachers normally deal with texts in the foreign language classroom (cf. 4.8). We wanted to make sure that the 'traditional' approach of the control programmes would not include the same type of systematic attention to the use of context realized in the training programme.

The control programmes of the pilot had not contained exercises, only reading texts. The teachers were told to deal with the texts in the classroom just as they would normally do. The only restriction was not to spend more than 40 minutes on each of the seven texts. During the pilot we noticed that, just as in the 1987-project, there was a certain overlap between the contents of the training and control conditions. The lessons of the control condition sometimes included aspects that might be considered to be context use. Some of the teachers who executed the control reading programmes briefly introduced text-topic (general context) or pointed at the loan words and cognates of a reading text (local context). Measures were taken to avoid that the lessons of the control condition of the experiment, which took place one year later, would include such aspects of context use. The tasks of the control programmes were based on our classrooms observation in the pilot study and included reading text parts aloud, answering questions and translation exercises. They did not embrace pre-reading activities, nor tasks in which students had to deal with loan words and cognates (cf. 4.8).

These measures were not sufficient. During our classroom observations we noticed that the two teachers of the control condition, and especially the teacher for Spanish, regularly drew the students' attention to the loan words and cognates in reading texts. This was something that formed part of their standard procedure of explaining unknown words. Although this type of context strategy is not explicitly referred to in our training programmes, recognizing cognates and loan words can be considered to be context use as well (Van Esch & Pérez Ruiz, 1994). As explained in Chapters 2 and 3, word-recognition ability is of major importance in information processing.

This means that the control programmes in this follow-up contained exercises that were probably also important for the training of the use of context as part of reading comprehension. This is confirmed by one of our research results. The majority of the students of control class CS1 (Spanish) believed that the control programme had taught them how to guess the meaning of unknown words.

Planning

There are two aspects of our planning that need some further examination: our attempts to integrate the experiment in the everyday practice of foreign language teaching at the two schools involved in our project, the duration of our training programme and the time-tables of the school-classes.

The training and control programmes had not become fully integrated in the foreign language curriculum. In Chapter 4, we explained that for reasons of ecological

validity, this integration was considered to be important. We had adjusted our materials to the language proficiency level of the students and part of the training programme was aimed at making students aware of the fact that the techniques learned could also be used outside the lessons of the training programme. This transfer of context use to other classroom activities was not easy to attain during the experiment. For practical reasons - the students had to practise commercial correspondence for their examinations later that year - there were two separate course contents during the execution of our experiment, i.e. our training and commercial correspondence. Consequently, the students experienced the training programme as being something extra, which had little to do with their regular classroom activities and examinations.

A second point of discussion is our choice for an intensive seven-week training programme instead of a programme over a larger period of time. A short training period means that the students are given little time to reflect on and really absorb the context strategies learnt. It is not unthinkable that quantitative effects of context-use training are not yet manifest after seven weeks, but only after a certain period of time. This observation is not necessarily in contradiction with the results of the earlier project (Van Esch, 1987), which found positive effects of training the use of context for beginning students of Spanish with the same seven week training programme as used in this follow-up. As explained above, there were reasons to assume that what had been interpreted as significant positive influence of context training, could (also) be contributed to the positive effects of the availability of explicit help from a teacher. Therefore, neither the results of the 1987-project, nor of this follow-up can conclusively answer the question whether it would have been better to spread the lessons of our training over a larger period of time.

Drop out

There was only a limited group of students at the two schools in Germany that followed the Spanish or English language courses. We started our investigation with a number of 153 subjects. The choice for a pre-test post-test control group design (cf. 4.4) implied that the data of each subject that had not performed one of the pre- or post-tests could not be used for the final analyses of the data. In our case, this meant that only 92 subjects remained, in the end. There were 25 subjects left of the Spanish training condition, 24 of the Spanish control condition, 25 of the English training condition and 18 of the English control condition. Therefore, the groups that could be compared with each other were rather small. For the analyses at the class level, we even had to decide not to take into account the results of training class TE3 (English) because only four subjects were left. Because the foreign language lessons were compulsory, there was no clear explanation for this drop out of subjects. Students may have been sick or may have had to skip lessons for other legitimate reasons.

Choice of instruments

The data gathered in this study were both quantitative (tests and questionnaires) and qualitative (teachers' interviews). The selection of the tests of our experiment was extensively discussed in 4.9. In that particular section, we saw that the reliability of most of the tests and questionnaires was established in advance, i.e. in the 1987-project and/or in our pilot study. The context tests and English vocabulary test were used in the main experiment without further adaptations. The Spanish vocabulary test needed to be adjusted. Finally, instead of two reading comprehension tests, 'cocktails' of multiple-choice cloze tests and a text with multiple-choice questions were used. Regarding the latter tests, the reliability could not be established in the pilot experiment: the reading tests for Spanish would not be offered to the students with German translations of the multiple-choice options, such as had been the case in the pilot, and two of the English multiple-choice cloze tests had not been tried out during the pilot. A pilot study carried out by CITO, however, had demonstrated that the four English multiple-choice cloze tests used in our experiment were reliable instruments for Dutch pupils at the MAVO-level (Van Zuylen, 1991).

Of course, we checked whether the execution of our tests during the main experiment had been satisfactory. The statistics of the main experiment can be found in Appendix XVIII. In the case of the Spanish and English context tests, we found reliability coefficients that were higher than those found in the pilot (Spanish: reliability coefficient context pre-test = .91, reliability coefficient context post-test = .84; English: reliability coefficient context pre-test = .78, reliability coefficient context post-test = .72). The reliability coefficients of the reading comprehension tests were somewhat lower than of the context tests (Spanish: reliability coefficient reading comprehension pre-test = .72, reliability coefficient reading comprehension post-test = .81; English: reliability coefficient reading comprehension pre-test = .69, reliability coefficient reading comprehension post-test = .68). These results show that our assumption that the larger number of items of the reading comprehension 'cocktails' (35 items instead of 20-25 in the pilot) should positively affect the reliability of the reading comprehension tests (cf. 4.9.3), could not always be confirmed. An explanation for the low reliability coefficients of the reading comprehension tests may be that the subjects in Germany were not used to perform multiple-choice tests. In Germany, it is not allowed to apply multiple-choice tests for official examinations. Although during the preparatory stage of the main experiment, the students practised multiple-choice tests (cf. 4.6), this preparation may not have been sufficient.

Below we will discuss the following aspects of our choice of instruments: the use of open context tests, the use of multiple-choice reading comprehension tests, our decision to use different tests in the pre- and post-tests sessions and the instruments we used to measure the students' and teachers' opinions about our programmes and tests.

The reason to decide to use open context tests was that one of the conclusions of the 1987-project was that this type of test is a valid instrument to measure context use. A disadvantage of using an open test, however, is that the rating may be subjective. To avoid this, open tests are often corrected by more than one rater. In that case, the interrater reliability is established and is considered to be a measure of the reliability of the correction of the test. Another solution is to establish in advance what answers will be correct and what answers not. As outlined in Chapter 5, we opted for this second solution. In our case, this implied that the tests were corrected three times, strictly according to previously established protocols (cf. 5.1.1).

For reasons of criterion validity, we decided to use other instruments for text comprehension in our investigation than in the 1987-project (cf. 4.2). In the earlier experiment open cloze tests were used, where in this follow-up project we used multiple-choice cloze tests and texts with multiple-choice questions. The two tests focus on different aspects of reading comprehension: in texts with multiple-choice questions both global and detailed comprehension is tested, whereas a multiple-choice cloze test measures comprehension of the main arguments in a text. At the time of the planning and elaboration of this follow-up our choice for a 'cocktail' of both types of multiple-choice reading tests seemed to be a suitable alternative for the open cloze. The first type of test - texts with multiple-choice questions - already formed part of standard reading comprehension tests developed by CITO for the final examinations of secondary education in the Netherlands. Also the multiple-choice cloze test appeared to be a valid instrument to measure reading comprehension (cf. 4.9.3). Nowadays, both types of reading tests are included in the standard examinations of secondary education in the Netherlands.

Besides criterion-validity, there were two other reasons to opt for multiple-choice tests. Firstly, the results of a multiple-choice tests can be easily and rapidly established. Secondly, it is established in advance what answers are correct. In other words, the interrater reliability of a multiple-choice test is not an issue. In the case of open tests, such as the context tests referred to above and the cloze tests of the 1987-project, the interrater reliability has to be established afterwards or other measures have to be taken to be sure that the correction of the test is reliable.

Our decision to use multiple-choice reading tests, however, may also have had disadvantages. Firstly, it is quite likely that a multiple-choice test is less sensitive than an open test, such as used in the 1987-project. This could partially explain why none of the school classes of the experiment showed a significant improvement in reading comprehension. Secondly, as explained above, the subjects of our experiment were unfamiliar with multiple-choice tests.

In 4.9 we explained why we decided to use different tests for the pre- and post-tests sessions. Regarding the context tests, this was a prerequisite, because words that had been 'guessed' correctly during a pre-test were, from that moment on, potentially familiar to the readers. Another argument was that in a normal school situation students would not have to make the same test twice within a relatively short time. Consequently, also the reading comprehension pre-test was different from the post-

test.

The tests of the post-test sessions were slightly more difficult than the pre-tests (cf. Appendix XVIII). This, however, can only partly explain why, except for the performance on context use of the Spanish training group, all overall mean gain scores were negative. A more likely explanation is that, in the course of the experiment, the students got used to the lessons of programmes and were, therefore, less motivated to perform the tests, than at the beginning of the experiment. In the case of Spanish, it is very likely that motivational factors played a part.

A last point of discussion are the instruments we used to measure the students' and teachers' evaluations of our programmes and tests.

After the experiment, we asked the students to evaluate the programmes and tests. The results of this evaluation show some apparent inconsistencies with the test results. Especially the results for English are difficult to interpret. According to their own opinion, the students of class TE2 and TE3 benefited from the training programme. This, however, could not be confirmed by significant better test results on context use and reading comprehension. An explanation for these results may be that students are not very good at judging themselves. A second explanation may be that a strong performance on the tests does not always combine with the students' feelings about the training programme, and vice versa.

The outcomes of the teachers' interviews provided much information about the execution of the experiment. We have to be very careful, however, to draw conclusions from qualitative results that cannot be confirmed by quantitative outcomes. An open interview is not always a reliable instrument. The fact that the teachers were interviewed after the experiment, when they had already been informed about the lack of results, may have influenced their opinions about programmes, tests and execution.

7.2.2. Reader-dependent variables

In this sub-section we will discuss the results related to our third research question. This question dealt with the importance of vocabulary knowledge, attitude towards the foreign language, self-confidence and the student's own estimate of his or her context use. We used vocabulary tests and a set of questionnaires to measure these variables. In Chapter 4, the reliability of these instruments was justified. The vocabulary tests were tried out during the pre-tests (cf. 4.9.1). The questionnaires were also used in the 1987-project (cf. 4.9.4).

Our follow-up confirmed the importance of vocabulary knowledge (Esch, 1987). For Spanish, the low correlations of vocabulary knowledge and the reading comprehension test post-test can be partly explained by the fact that the results of class TS2 in this test were not reliable.

According to Spoelders (1978) a positive attitude is more relevant for beginning foreign language students than for advanced students. This explains why we found significant correlations of the factor 'attitude' for Spanish, but not for English.

Considering the fact that the overall mean score on attitude was relatively high, i.e. 57.36 out of 90 (maximum total score), a competing, more tentative explanation could be that both weak and strong students had a positive attitude towards this foreign language, due to a general agreement nowadays about the social need to be able to speak and understand English.

There was a discrepancy between the students' own estimate of their context use and their actual performance on the context and reading comprehension tests. This is not really surprising, because a comparable result was found in the 1987-project. In this study, the results of TE1 and TE2 discussed above showed a more or less comparable inconsistency. Two possible explanations could be that the students simply were not good at evaluating themselves or that using context strategies did not guarantee a high score on our context and reading comprehension tests.

Finally, a tentative explanation for the significant correlation between 'self-confidence' and the Spanish context post-test could be that the weaker students, i.e. the ones who followed the training programme in the use of context, scored low on self-confidence after the training programme because they had lost faith in their ability to deal with Spanish texts.

7.3. Implications for research on and teaching of non-native reading

The experiment described in this study has not led to clear results. This can partly be explained by the conditions under which we executed our experiment, partly by the complex of processes that determine whether a reading training in the foreign language classroom will be successful.

The conditions under which we executed the experiment were not always as we had meant them to be. In the case of Spanish, this even led to the situation that no reliable comparison could be made between training and control groups. On the whole, our training programmes in the use of context did not lead to better results than the control reading programmes. It is not unlikely that the training and control programmes of our experiment were both important for the training of context use and reading comprehension. The outcomes of Mulder's study (Mulder, 1996), discussed in Chapter 3, support this assumption. She found that both the training of reading strategies and 'common practice' substantially improved the subjects' reading comprehension ability. In the previous section, we saw that also the control programmes contained elements of context-use training, such as analysing loan words and cognates. Besides, in the lessons of the control condition, the students did not have to guess the meaning of unknown words, but these meaning were given by the teachers. This probably helped them in the process of answering questions about text contents and translating text fragments.

A positive finding of our study was that all teachers believed that their students had become more self-reliant after the training programme. This at least suggests that the students gained from the training programmes with regard to metacognitive skills - e.g. monitoring reading tasks and solving reading problems by themselves -, although this is not confirmed by quantitative outcomes. Despite the fact that the results of our training programmes in the use of context were limited and we experienced many difficulties, the results of this study may contribute to a better understanding of factors that influence the success of training reading in the foreign language classroom.

The last 15 years or so, researchers have speculated about the effectiveness of strategy training and have tried to find empirical proof for their assumptions. Points of discussion have been objectives, effects and the limiting conditions of teaching reading strategies in the foreign language classroom. In Chapter 3, a review was given of studies in which the effects of reading training were assessed. We saw that positive effects of strategy training were found, as well as no effects or contradictory results. We also referred to theories about the possible existence of a 'threshold' of linguistic proficiency below which it would be difficult to successfully apply reading strategies (Clarke, 1980; Alderson, 1984; Bossers, 1991). The discussion between Bossers (1993) and Westhoff (1993), reviewed in Chapter 3, is a good example of different ways of approaching the teaching of non-native reading. According to Bossers (1993), it is not necessary to teach reading strategies in the foreign language classroom, because once students reach a critical level of foreign language proficiency, they will be able to transfer their first language reading skills to a non-native reading setting. Instead, it would be more useful to dedicate attention to the expansion of the students' foreign language vocabulary, the development of their decoding skills and other linguistic aspects of reading comprehension. As outlined in Chapter 3, both linguistic and strategic aspects of non-native reading are important. Although we recognized the importance of linguistic proficiency and the existence of 'short-circuits' (Clarke, 1980) in the process of reading comprehension, we also assumed that it would be possible to override the effects of low non-native proficiency by activating appropriate content schemata and using other sources of contextual information. This was based on Stanovich's theory of compensatory processing in first language reading (Stanovich, 1991), outlined in Chapter 2. In first language reading schema-theoric, text-based and compensatory models of reading have to be considered as complementary views on reading and not as competing views. According to our view, this model also adequately explains non-native reading (cf. Chapter 3). Our main argument to conduct our training programme in the use of context was that linguistic knowledge alone, would not be sufficient to achieve reading comprehension, but that also metacognitive skills were indispensable (Stanovich, 1991, Bernhardt, 1991).

In our experiment we deliberately manipulated the variable non-native proficiency. The students for Spanish were beginners, whereas the subjects involved in the English training and control programmes were intermediate students. On the basis of

our quantitative research findings we could not establish whether the Spanish or the English training programme had been more successful. It was clear, however, that the execution of the Spanish training programme led to more problems than the execution of the English programme. An explanation for this could be the low proficiency level of the students that performed the Spanish training programme. This would suggest that language proficiency did play a part in our investigation, after all. It is not unlikely that the low language proficiency level of the Spanish students inhibited the successful application of the context strategies acquired and led to a situation in which both teachers and students did not feel comfortable. Probably the students knew what they were expected to do (infer the meaning of unknown words), even knew how they could solve this (by applying the context strategies acquired) but simply had not sufficient tools (linguistic proficiency) to successfully accomplish this task. In other words, they were likely to possess sufficient metacognitive knowledge, but were probably not always able to apply this knowledge in their metacognitive processing. In an update on the current research of cognitive strategy-use instruction for first language learners in the United States, Pressley et al.(1995) questions the effectiveness of short interventions like our training programme, because they mainly focus on strategic aspects and not on content aspects like expansion of vocabulary-size. This is probably equally important in a non-native setting. On the basis of the results of this study, however, we cannot conclusively answer to what degree the low proficiency level of the Spanish training group impeded the use of context. Although the two teachers mentioned lack of language proficiency as the main reason for the disappointing outcomes, there are results of our project that seem to conflict with this explanation: despite the low language proficiency level, the mean gain scores of TS1 were positive, both in the case of context use and reading comprehension. It is likely, though, that the low proficiency level of the students of the Spanish training condition interfered with what was learnt about context use and that this led to frustrations. However, not only linguistic proficiency was an important variable in our experiment, but also other factors played a part.

In Chapter 3 we discussed Bernhardt's model of non-native reading (Bernhardt, 1991). Bernhardt suggests a whole-process approach towards reading, in which prior knowledge and (cognitive and metacognitive) strategy use interact with linguistic proficiency. In this model, linguistic skills are essential, but not sufficient. The model of non-native reading explains how language components, strategy use and background knowledge interact at different levels of non-native proficiency. According to Bernard's model, metacognitive processing and background knowledge are relevant factors in both advanced and beginning non-native reading. Our research results confirm Bernhardt's model: our subjects dealt with the input of the teaching in different ways and the effects of our teaching could be related to a wide array of factors: linguistic skills, personal variables such as motivation, and situational factors such as teacher, teaching materials, examination exigencies and time-table.

What insights does this study contribute into how reading comprehension has to be dealt with in the classroom?

One of the main findings of our research is that the value of our training programme in the use of context could not be established by means of overall effects. In our case, we found differences between school classes and even between individual students. This means that our assumption that the same input (a training programme in the use of context) will lead to the same output (improves reading comprehension) was not correct. There is variation in the students' cognitive and metacognitive abilities, linguistic proficiency, background knowledge, attitude towards the language, self-confidence, intelligence and preference of approaching a text (cf. Chapter 3). Current views on teaching recognize that there are differences in the learning processes of individual students. In the Netherlands, the newly introduced concept of schools as 'places of study' focuses on autonomous learning. This means that a teaching environment will be provided in which the broad personal development of all pupils is encouraged (cf. Chapter 1).

It is questionable whether context-use training should be organized in the same way as the training programmes in this study. First of all, the point made by Pressley (1995) that it may not be effective to organize a short intervention that mainly focuses on the strategic aspects of reading is important. Looking back at the experiment reported in this study, there are reasons to believe that it would be better to spread context-use training over a larger period of time than was possible in our experiment and to integrate metacognitive strategy training in the foreign language (reading) curriculum. This probably avoids that students, and especially beginning foreign language learners, lose their motivation. We do not believe that context-use training is not suitable for beginning learners: in our view, it is crucial to dedicate attention to the development of strategic skills, also at an elementary level of non-native proficiency. This type of instruction, however should be alternated with exercises that extend the students' vocabulary and develop their bottom-up skills. Unfortunately, in our project it had not been possible to spread the lessons of the programmes over a larger period for practical reasons.

More or less similar ideas are reflected in Bernhardt's study (1991), who asserts that students will have to work on expansion of vocabulary as well as on metacognitive insights. Vocabulary should not be presented in the form of word lists, but within a broader context. She underlines that teachers have to recognize that students already possess metacognitive skills and should work from "how readers understand the texts they are asked to read and should help students to reformulate their understandings when misunderstandings occur" (p 42). Bernhardt postulates that training word-recognition skills is probably only necessary for poor non-native readers and readers that have to learn a different graphic system (e.g. Russian, Chinese). Segalowitz et al (1991) assert that, if word-recognition training is required, three criteria have to be met. Firstly, teachers have to focus on the appropriate processing component. Training word recognition in a language with a

graphic system similar to the mother tongue requires a different type of word-recognition training than a language with a totally different graphic system. Secondly, the training has to be consistent and repetitive. To attain automatization of word-recognition processes it is essential that a student should often process the same material (words, letter strings) and each time in the same way. Thirdly, the training situation should focus on perceptual and cognitive demands that a reader would also be confronted with in a real life situation. Finally, the model developed by Bernhardt implies that reading instruction has to be flexible. It has to meet students' needs at particular stages of their reading comprehension ability and should accommodate to the students' needs over time.

In this type of instruction, the role of the teacher is important. Teachers will not only have to improve their students' linguistic skills, but will also have to make students aware of their metacognitive abilities and individually guide them through the process of improving their metacognitive skills. This is not an easy task. This is confirmed by current experiences in the Netherlands with the implementation of the school as 'a place of study'. Teachers have to adopt a totally different attitude towards the learning of their pupils. It is difficult to find a method that meets the needs of each individual student. We do believe that context-use training has the potential to meet needs of different students at different points in their process of learning to deal with non-native texts. The results of such a classroom intervention, however, may be different for each individual student.

In our experiment, we have seen some examples of the important role of the teacher. Both characteristics of individual teachers and their previous experience with context-use training have appeared to be important variables. Teacher TS1, for example, was an experienced context-use trainer and a person who was very concerned with the well-being of his students. The combination of these two characteristics probably contributed to the relatively good performance of TS1. It is, therefore, important that not only students, but also teachers should be trained to deal with the principles of context-use training.

7.4. Suggestions for further research

Considering the complexity of reading comprehension, it is necessary to focus on the processes underlying reading comprehension. In our study we found that for some individual students our training programme had been very successful. It would be useful to find out how these individual students dealt with our training programme and what factors contributed to these positive outcomes. Additionally, it would be helpful to find answers to questions such as: How do individual readers deal with a text? How do they monitor their reading? What classroom interactions contribute to improved reading comprehension ability? What kind of interactions take place in a classroom setting? How can we help individual students to improve their comprehension skills? As explained in Chapter 1, current views on teaching reading in the classroom are more concerned with differences in individual learning

processes of students and try to meet individual needs. The type of non-native reading research suggested above, could supply the necessary information and insights to deal with this new situation.

Another point to take into consideration in further research is the assessment of tasks that teach students to improve their metacognitive ability. In our study, we used a survey with context 'rules' to guide the students' process. In other studies different types of tasks are used. It would be interesting to find out, whether certain types of tasks lead to more improvements in metacognitive ability than others.

There is a need to investigate the role of native language reading skills in non-native reading. According to Bossers (1993), it is not necessary to teach strategy use in non-native reading because once students reach the so-called 'threshold level' of non-native proficiency, they will transfer their first language reading skills to a non-native reading setting. As outlined in this study, we believe that it is also necessary to dedicate attention to strategy use in the foreign language classroom. This is confirmed by recent findings. The results of a study executed by Bimmel (1999), demonstrate that there is no spontaneous transfer of native language reading skills to non-native reading. In our study, we did not compare the subjects' foreign language reading skills with their native reading behaviour. Consequently, we could not establish whether and/or to what degree native reading skills were transferred.

Reading research should not only focus on the transfer of native reading skills to non-native reading, but also on the transfer of skilled reading in one particular non-native language to reading in another foreign language. For practical reasons, we decided to drop the fourth research question, i.e. whether there will be a transfer effect of a training programme in context-use in one foreign language to the performance on context use and reading comprehension in another foreign language. Follow-up research is necessary to find an answer to this and similar questions.

As pointed out in this study, vocabulary is considered to be critical in non-native reading. The correlations found in our study, however, were not high. Research on the relationship between vocabulary and foreign language reading comprehension ability is relatively slight and not always convincing (cf. Barnett, 1989). It is important that future research explores the role of vocabulary in non-native reading.

At the beginning of this work, we explained that it would be interesting to assess whether there are differences between the effects of a context-use training programme at an elementary and at an intermediate non-native proficiency level. On the basis of the quantitative results of this study, this question could not be answered. Therefore, the assumed relationship between level of linguistic proficiency and the success of strategy-use training needs to be investigated further.

References

- Aarnoutse, C.A.J. (1982). *Aspecten van begripend lezen in het vierde leerjaar van het gewoon lager onderwijs*. Doctoral dissertation University of Nijmegen (KUN).
- Adams, S.J. (1982). Scripts and the recognition of unfamiliar vocabulary: Enhancing second language reading skills. *Modern Language Journal*, 66 (2), 155-159.
- Alderson, J.C. (1979). Scoring procedures for use on cloze-tests. In C.A. Yorio, K. Perkins & J. Schachter (Eds.), *On TESOL 1979. The learner in focus* (pp. 193-205). Washington D.C: TESOL.
- Alderson, J.C. (1984). Reading in a foreign language: A reading problem or a language problem? In J.C. Alderson & A.H. Urquhart (Eds.), *Reading in a foreign language* (pp. 1-24). London: Longman.
- Alderson, J.C. & Urquhart, A.H. (1985/1988). This test is unfair: I'm not an economist. In P.L. Carrell, J. Devine & D.E. Eskey (Eds.), *Interactive approaches to second language reading* (pp. 168-182). Cambridge: Cambridge University Press. [Reprinted from P. Hauptman, R. Leblanc & M. Bingham Wesche (eds.), *Second language performance testing*. Ottawa: University of Ottawa Press, 1985].
- Anderson, J. (1976). *Psycholinguistic experiments in foreign language testing*. Queensland: University of Queensland Press.
- Anderson, R.C. & Pearson, P.D. (1984/1988). A schema-theoretic view of basic processes in reading comprehension. In P.L. Carrell, J. Devine & D.E. Eskey (Eds.), *Interactive approaches to second language reading* (pp. 37-55). [Reprinted from P.D. Pearson (Ed.), *Handbook of reading research*. New York: Longman, 1984].
- Barnett, M. A. (1986). Syntactic and lexical/semantic skill in foreign language reading: Importance and interaction. *Modern Language Journal*, 70, 343-349.
- Barnett, M. A. (1989). *More than meets the eye. Foreign language reading*. Englewood Cliffs, New Jersey: Prentice Hall.
- Bartlett, F.C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press.
- Bensoussan, M. & Laufer, B. (1984). Lexical guessing in context in EFL comprehension. *Journal of Research in Reading*, 7, 15-32.
- Bernhardt, E.B. (1987). Cognitive processes in L2: An examination of reading behaviors. In J. Lantolf & A. Labarca (Eds.), *Research on second language acquisition in classroom settings: Delaware symposium 6* (pp. 35-50). Norwood, N.J.: Ablex.

- Bernhardt, E.B. (1991). A psycholinguistic perspective on second language literacy. In J.H. Hulstijn & J.F. Matter (Eds.), *Reading in two languages* (pp.31-44). AILA Review (Vol. 8). Amsterdam: Free University Press.
- Bimmel, P.E. (1999). *Training en transfer van leesstrategieën. Training in de moedertaal en transfer naar een vreemde taal. Een effectstudie bij leerlingen uit het voortgezet onderwijs*. Doctoral dissertation University of Utrecht (RUU). 's-Hertogenbosch: Malmberg.
- Block, E.L. (1986). The comprehension strategies of second language readers. *TESOL Quarterly*, 20 (3), 463-494.
- Boekaerts, M. & Simons, P.R.-J. (1993). *Leren en instructie. Psychologie van de leerling en het leerproces*. Assen: Dekker & Van de Vegt.
- Bormuth, J.R. (1967). Optimum sample size and cloze test length in readability measurement. *Journal of Educational Measurement*, 2, 111-116.
- Bossers, B. (1989). Lezen in de tweede taal: Een taal of leesprobleem? *Toegepaste Taalwetenschap in Artikelen*, 38, 176-188.
- Bossers, B. (1991). On thresholds, ceilings and short-circuits: the relation between L1 reading, L2 reading and L2 knowledge. In Hulstijn, J.H. & J.F. Matter (Eds.), *Reading in two languages* (pp. 45-60). AILA Review (Vol. 8). Amsterdam: Free University Press.
- Bossers, B. (1992). *Reading in two languages: A study of reading comprehension in Dutch as a second foreign language and in Turkish as a first language*. Doctoral dissertation University Amsterdam (VU). Rotterdam: Van Driel.
- Bossers, B. (1993). VT-lezen: Leesstrategieën of taalkennis? *Levende Talen*, 478, 133-139.
- Brown, A.L. (1980). Metacognitive development and reading. In R.J. Spiro, B.C. Bruce & W.F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp.453-481). Hillsdale, N.J.: Erlbaum.
- Brown, A. L. & Palinscar, A.S. (1982). Inducing strategic learning from texts by means of informed self-control training. *Topics in learning and learning disabilities*, (Vol. 2), Special issue on metacognition and learning disabilities, 1-17.
- Brügel, H. (1989). *El español de hoy*. Wolfenbüttel: Heckners Verlag.
- Carpay, J.A.M. (1975). *Onderwijsleerpsychologie en leergangontwikkeling in het moderne-vreemdetalenonderwijs*. Groningen: Tjeenk Willink.
- Carrell, P.L. (1981). Culture-specific schemata in L2 comprehension. In R. Orem & J. Haskell (Eds.), *Selected papers from the ninth Illinois TESOL/BE annual convention, the first Midwest TESOL conference* (pp.123-132). Chicago: Illinois TESOL/BE.
- Carrell, P.L. (1983). Some issues in studying the role of schemata, or background knowledge, in second language comprehension. *Reading in a Foreign Language*, 1 (2), 81-112.
- Carrell, P.L. (1988a). Introduction. In P.L. Carrell, J. Devine & D.E. Eskey (Eds.), *Interactive approaches to second language reading* (pp. 1-7). Cambridge: Cambridge University Press.

- Carrell, P.L. (1988b). Some causes of text-boundedness and schema interference in ESL reading. In P.L. Carrell, J. Devine & D.E. Eskey (Eds.), *Interactive approaches to second language reading* (pp. 101-113). Cambridge: Cambridge University Press.
- Carrell, P.L. (1989). Metacognitive awareness and second language reading. *Modern Language Journal*, 73, 121-134.
- Carrell, P.L. (1991). Second language reading: Reading ability or language proficiency? *Applied Linguistics*, 12, 159-179.
- Carrell, P.L. & Eisterhold, J.C. (1983/1988). Schema-theory and ESL reading pedagogy. In P.L. Carrell, J. Devine & D.E. Eskey (Eds.), *Interactive approaches to second language reading* (pp. 73-92). Cambridge: Cambridge University Press. [Reprinted from *TESOL Quarterly*, 17, (4), 1983, 553-573].
- Carrell, P.L., J. Devine & D.E. Eskey (Eds.) (1988). *Interactive approaches to second language reading*. Cambridge: Cambridge University Press.
- Carver, R. (1977-78). Toward a theory of reading comprehension and rauding. *Reading Research Quarterly*, 13, 8-64.
- Cattell, J. M. (1886). The time it takes to see and name objects. *Mind*, 11, 63-65.
- Claessen, J., Van Galen, A. & Oud-de Glas, M. (1978). *De behoeften aan moderne vreemde talen. Een onderzoek onder leerlingen, oud-leerlingen en scholen*. Nijmegen: Institute for Applied Sociology (ITS).
- Clarke, M.A. (1980/1988). The short circuit hypothesis of ESL reading - or when language competence interferes with reading performance. In P.L. Carrell, J. Devine & D.E. Eskey (Eds.), *Interactive approaches to second language reading* (pp. 114-124). Cambridge: Cambridge University Press. [Reprinted from *Modern Language Journal*, 64 (2), 1980, 203-209].
- Collins, A.M. & Quillian, M.R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 8, 240-247.
- Cummins, J. (1979). Cognitive academic language proficiency, linguistic interdependence, the optimum age question and some other matters. *Working Papers on Bilingualism*, 19, 197-205.
- CVE-MVT (1992). *Eindrapport*. Arnhem: Commissie Vernieuwing Eindexamens Moderne Vreemde Talen.
- Cziko, G. (1978). Differences in first and second language reading: The use of syntactic and discourse constraints. *Canadian Modern Language Review*, 34, 473-489.
- Cziko, G.A. (1980). Language competence and reading strategies: A comparison of first- and second-language oral reading errors. *Language Learning*, 30, 101-114.
- Devine, J. (1988a). The relationship between general language competence and second language reading proficiency: Implications for teaching. In P.L. Carrell, J. Devine & D.E. Eskey (Eds.), *Interactive approaches to second language reading* (pp.260-277). Cambridge: Cambridge University Press.
- Devine, J. (1988b). A case study of two readers: Models of reading and reading performance. In P.L. Carrell, J. Devine & D.E. Eskey (Eds.), *Interactive approaches to second language reading* (pp.127-139). Cambridge: Cambridge University Press.

- Eskey, D.E. (1988). Holding in the bottom: An interactive approach to the language problems of second language readers. In P.L. Carrell, J. Devine & D.E. Eskey (Eds.), *Interactive approaches to second language reading* (pp. 93-100). Cambridge: Cambridge University Press.
- Ferguson, G.A. & Takane, Y. (1989). *Statistical analysis in psychology and education* (6th ed.). New York: McGraw-Hill Book Company.
- Flavell, J.H. (1976). Metacognitive aspects of problem solving. In L.B. Resnick (Ed.), *The nature of intelligence* (pp. 231-236). Hillsdale, N.J.: Erlbaum.
- Floyd, P. & Carrell, P.L. (1987). Effects on ESL reading of teaching cultural content schemata. *Language Learning*, 37, 89-108.
- Fodor, J.A. (1983). *The modularity of mind*. Cambridge, Mass.: MIT Press.
- Goodman, K.S. (1967). Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist* 6, 126-135.
- Goodman, K.S. (1973). Psycholinguistic universals in the reading process. In F. Smith (Ed.), *Psycholinguistics and Reading* (pp. 21-28). New York: Holt, Rinehart & Winston.
- Goodman, K.S. (1981). Letter to the editor. *Reading Research Quarterly*, 16 (3), 477-478.
- Gough, P.B. (1972). One second of reading. In J.F. Kavanagh & J. G. Mattingley (Eds.), *Language by ear and by eye* (pp. 331-358). Cambridge, Mass.: The MIT Press.
- Grabe, W. (1991). Current developments in second language reading research. *TESOL Quarterly*, 25 (3), 375-406.
- Grabe, W. (1993). *Rethinking reading strategies: Research and instruction*. Paper presented at TESOL 93, Atlanta.
- Hacquebord, H. (1989). *Tekstbegrip van Turkse en Nederlandse leerlingen in het voortgezet onderwijs*. Studies over taalgebruik (Vol. 6). Doctoral dissertation University of Groningen (RUG). Dordrecht: Foris Publications.
- Håkanson, U., Masoliver, J. & Sandström, G. et al. (1990). *¡ Eso sí! 1*. Text book & workbook. (German edition: C. Segoviano, S. Segoviano-Rosenblum & C. Palaoro). Stuttgart: Ernst Klett Verlag.
- Hallf, H.M., Orthony, A. & Anderson, R.C. (1976). A context-sensitive representation of word meaning. *Memory and Cognition*, 4, 378-383.
- Halliday, M.A.K., & Hassan, R. (1976). *Cohesion in English*. London: Longman
- Haynes, M. (1993). Patterns and perils of guessing in second language reading. In T. Huckin, M. Haynes & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 46-62). Norwood, N.J.: Ablex.
- Hinofotis, T.A.B. (1976). *An investigation of the current validity of cloze testing as a measure of overall proficiency in English as a second language*. Unpublished doctoral dissertation, South Illinois University.
- Holley, F. (1973). A study of vocabulary learning in context: The effect of new-word density in German reading materials. *Foreign Language Annals*, 6, 339-347.

- Hosenfeld, C. (1977). A preliminary investigation of the reading strategies of successful and unsuccessful second language learners. *System*, 5 (2), 110-23.
- Hosenfeld, C. Arnold, V., Kirchofer, J., Laciura, J. & Wilson, L. (1981). Second language reading: A curricular sequence for teaching reading strategies. *Foreign Language Annals*, 14 (5), 415-422.
- Huck, S.W., Courmier, W.H. & Bounds, JR., W.G. (1974). *Reading statistics and research*. New York: Harper & Row.
- Hudson, T. (1982/1988). The effects of induced schemata on the "short circuit" in L2 reading: Non-decoding factors in L2 reading performance. In P.L. Carrell, J. Devine & D.E. Eskey (Eds.), *Interactive approaches to second language reading* (pp. 183-205). Cambridge: Cambridge University Press. [Reprinted from *Language Learning*, 32 (1), 1982, 3-31].
- Hulstijn, J.H. & J.F. Matter (Eds.) (1991). *Reading in two languages*. AILA Review (Vol. 8). Amsterdam: Free University Press.
- Jansen, K., Jasper, H.-H. & Stockwell, T. (1986). *English im Berufsfeld Wirtschaft*. Köln: Stam-Verlag.
- Johnson, P. (1981). Effects on reading comprehension of language complexity and cultural background of a text. *TESOL Quarterly*, 15 (2), 169-181.
- Johnson, P. (1982). Effects on reading comprehension of building background knowledge. *TESOL Quarterly*, 16 (4), 503-516.
- Jongsma, E.R. (1970). *The cloze procedure: A survey of research*. Occasional papers in reading. Bloomington, Ind.: ERIC Documents 050893.
- Just, M.A., & Carpenter, P.A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87, 329-354.
- Just, M.A., & Carpenter, P.A. (1987). *The psychology of reading and language comprehension*. Boston: Allyn & Bacon.
- Kern, R. (1989). Second language reading strategy instruction: Its effects on comprehension and word inference ability. *Modern Language Journal*, 73 (2), 135-149.
- Kintsch, W. & Van Dijk, T. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85 (5), 363-394.
- Klein-Braley, C. (1981). *Empirical investigations of cloze-tests. An examination of the validity of cloze-tests as tests of general language proficiency in English for German University Students*. Doctoral dissertation University of Duisburg.
- Klein-Braley, C. (1982). On the suitability of cloze tests as measures of reading comprehension. *Toegepaste Taalwetenschap in Artikelen*, 13, 49-61.
- LaBerge, D., & Samuels, S.J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 6, 293-323.
- Lee, J.F. (1986). Background knowledge & L2 reading. *Modern Language Journal*, 71 (1), 50-54.
- Lee, J.F. (1991). On the dual nature of the second-language reading proficiency of beginning language learners. In Teschner, R. V. (Ed.), *Assessing foreign language proficiency of undergraduates* (pp.198-215). Boston: Heinle.
- LeLoup, J.W. (1993). *The effect of disposition towards text topic on second language reading comprehension*. Paper presented at the annual conference of the American Association of Applied Linguistics (AAAL), Atlanta.

- Levine, M.G. & Haus, G J (1985) The effect of background knowledge on the reading comprehension of second language learners. *Foreign Language Annals*, 18 (5), 391-397.
- Macfarlane, M. (1987). *Focus in Success*. Berlin: Cornelsen & Oxford University Press GmbH.
- Mandell, P B. (1993). *Word recognition and comprehension* Paper presented at the annual conference of the American Association of Applied Linguistics (AAAL), Atlanta.
- Meijden, H. van der (1984). *Effecten van zelfvertrouwen en feedback op leesvaardigheid in een vreemde taal*. Unpublished master's thesis Department of Spanish University of Nijmegen (KUN).
- Mets, J. & Glas, C.A.W. (1988). *Een meerkeuze-invultoets ('cloze') geschikt voor eindexamen tekstbegrip?* Arnhem: Dutch National Institute for Educational Measurement (CITO). Interne documentatie nr. 289.
- Mondria, J.A. (1996). *Vocabulaireverwerving in het vreemde-talenonderwijs: De effecten van context raden op de retentie*. Doctoral dissertation University of Groningen (RUG). Groningen: Universiteitsdrukkerij.
- Mulder, H. (1996) *Training in leesstrategieën Vorm en rendement Een onderzoek naar het effect van vier trainingsvarianten op de leesvaardigheid Frans als vreemde taal* Doctoral dissertation Utrecht University (RUU). Amsterdam: Meulenhoff Educatief.
- Nagy, W. (1988). *Teaching vocabulary to improve reading comprehension*. Newark DE: International Reading Association.
- Nation, P. & Coady, J. (1988). Vocabulary and reading. In R. Carter & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 97-110). New York: Longman.
- Neuner, G. (1980). 'Lesen' und 'Verstehen' im kommunikativen Fremdsprachenunterricht. In H. Von Faber (Ed.), *Pariser Werkstattgespräche, 1978 Leseverstehen im Fremdsprachenunterricht* (pp. 98-111). München: Goethe Institut.
- Neuner, G., Krüger, M. & Grewer, U. (1981). *Übungstypologie zum kommunikativen Deutschunterricht*. Berlin: Langenscheidt.
- O&W (1986). *Examenprogramma moderne vreemde talen C/D-niveau in Ibo en mavo*. Den Haag: Ministerie van Onderwijs en Wetenschappen (Ministry of Education and Science).
- OC&W (1998). *Education in the Netherlands* [Online]. Available: <http://www.minocw.nl/english/edusyst/ednl00.htm> [1998, January 25]. Den Haag: Ministerie van Onderwijs, Cultuur & Wetenschappen (Ministry of Education, Culture and Science).
- Omaggio, A.C. (1979). Pictures and second language comprehension: Do they help? *Foreign Language Annals*, 12 (2), 107-116.
- Perfetti, C.A (1985). *Reading ability*, New York: Oxford.
- Perkins, K. & Bratten, S. (1983). The effects of word frequency and contextual richness on ESL students' word identification abilities. *Journal of Research in Reading*, 6 (2), 119-128.

- Pit, M. (1996). Voor taalmiddelen geen plaats? Het nieuwe niveaustelsel voor examens moderne vreemde talen. *Levende Talen*, 509, 229-231.
- Pit, M. & Zuylen, N. van (1996). Krenten in de pap. De ontwikkeling van examenopgaven voor gericht lezen in de moderne vreemde talen. *Levende Talen*, 508, 130-134.
- Potter, T.C. (1968). *A taxonomy of cloze research. Part I: Readability and reading comprehension*. Inglewood, Cal.: South-West Regional Laboratory for Educational Research and Development.
- Pressley, M., Woloshyn, V & Associates (1995). *Cognitive strategy instruction that really improves children's academic performance* (2nd. ed). Cambridge Ma.: Brookline Books.
- Rankin, E.F. (1965). Cloze procedure - a survey of research. In E.S. Thurston, & L.E. Hafner (Eds.), *The philosophical and sociological bases of reading. Fourteenth yearbook of the national reading conference* (pp. 133-150). Milwaukee: National Reading Conference.
- Rayner, K. & Pollatsek, A. (1989). *The psychology of reading*. Englewood Cliffs, N.J.: Prentice Hall.
- Rosenblatt, L. (1978). *The reader, the text and the poem: The transactional theory of the literary work*. Carbondale Ill.: Southern Illinois University Press.
- Ruddell, R.B. (1976). Reading comprehension and structural redundancy in written material. *Proceedings of the International Reading Association*, 10, 308-311.
- Rumelhart, D.E. (1977). Toward an interactive model of reading. In S. Dornic (Ed.), *Attention and performance (Vol. 6)* (pp.573-603). Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Rumelhart, D.E. & Orthony, A. (1977). The representation of knowledge in memory. In R.C. Anderson, R.J. Spiro & W.E. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp. 99-135). Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Rumelhart, D.E., McClelland, J.L. & the Ph.D. Research Group (1986). *Parallel distributed processing: Explorations in the microstructure of cognition (Vol. 1-2)*. Cambridge, Mass.: MIT Press.
- Samuels, S.J. & Kamil, M.L. (1984/1988). Models of the reading process. In P.L. Carrell, J. Devine & D.E. Eskey (Eds.), *Interactive approaches to second language reading* (pp. 22-36). Cambridge: Cambridge University Press. (pp. 22-36). [Reprinted from P.D. Pearson (Ed.), *Handbook of reading research*. New York: Longman, 1984].
- Saragi, T., Nation, I.S.P. & Meister, G.F. (1978). Vocabulary learning and reading. *System*, 6, 72-88.
- Sarig, G. (1987). High-level reading in the first and in the foreign language: Some comparative process data. In P.L. Carrell, J. Devine & D.E. Eskey (Eds.), *Research in reading in English as a second language* (pp. 107-120). Washington: TESOL.

- Saville-Troike, M. (1979). Reading and the audio-lingual method. In R. Mackay, B. Barkman & R.R. Jordan (Eds.), *Reading in a second language. Hypotheses, organization and practice* (pp.24-35). Rowley, Mass.: Newbury House.
- Schank, R.C. & Abelson, R.P. (1977). *Scripts, plans, goals and understanding*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Schouten-van Parreren, C. (1983). Wisseling van de wacht in de vreemde-talendidaktiek? De receptief-handelspsychologische benadering. *Levende Talen*, 378, 22-29.
- Schouten-van Parreren, C. (1985). *Woorden leren in het vreemde-talenonderwijs*. Doctoral dissertation Free University Amsterdam (VU). Apeldoorn: Van Walraven.
- Segalowitz, N., Poulsen, C., & Komeda, M. (1991). Lower level components or reading skill in higher level bilinguals: Implications for reading instruction. In Hulstijn & J.F. Matter (Eds.), *Reading in two languages* (pp.15-30). AILA Review (Vol. 8). Amsterdam: Free University Press.
- Smith, E.E. & Medin, D.C. (1981). *Categories and concepts*. Cambridge, Mass.: Harvard University Press.
- Smith, F. (1971). *Understanding reading: A psycholinguistic analysis of reading and learning to read*. New York: Holt, Rinehart and Winston.
- Spoelders, M. (1978). Taalattitudemeting in theorie en praktijk. *Toegepaste Taalwetenschap in Artikelen*, 5, 15-28.
- Staatsen, F., Meijer, D., Mulder, H. Stoks, G. & Toorenborg, H. van (1994). *Moderne vreemde talen in de basisvorming. Kerndoelen in de praktijk*. Bussum: Dick Coutinho.
- Stanovich, K.E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, 16, 32-71.
- Stanovich, K.E. (1984). The interactive compensatory model of reading: A confluence of developmental, experimental, and educational psychology. *Rase*, 35, 11-19.
- Stanovich, K.E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21, 360-407.
- Stanovich, K.E. (1991). Word recognition: Changing perspectives. In R. Barr, M. Kamil, P. Mosenthal & D. Pearson (Eds.), *Handbook of reading research* (Vol. 2) (pp. 418-452). New York: Longman.
- Stansfield, F. (1980). The cloze-procedure as a progress test. *Hispania*, 63, 715-718.
- Steffensen, M.S. & Joag-Dev, C. (1984). Cultural knowledge and reading. In J.C. Alderson & A.H. Urquhart (Eds.), *Reading in a foreign language* (pp.48-61). New York: Longman.
- Struik, C. (1996). Help, de commissie verzuipt. *Levende Talen*, 509, 232-233.
- Swaffar, J.K., Arens, K. & Byrnes, H. (1991). *Reading for meaning: An integrated approach to language learning*. Englewood Cliffs, N.J.: Prentice Hall.

- Taylor, W.L. (1953). "Cloze procedure": A new tool for measuring readability. *Journalism Quarterly*, 30, 415-433.
- Van Els, T., Extra, G., Os, Ch. van & Janssen-van Dieten, A. (1984). *Applied linguistics and the learning and teaching of foreign languages*. London: Edward Arnold.
- Van Esch, C. (1987). *Contextgebruik en begrijpend lezen in een vreemde taal. Evaluatie van een trainingsprogramma*. Doctoral dissertation University of Nijmegen (KUN). Groningen: Wolters-Noordhoff.
- Van Esch, K. & Pérez Ruiz, A. (1994). Cognates, woordenschat en lezen in een vreemde taal. *Levende Talen*, 487, 73-76.
- Van Krieken, R. (1984). *Tekstbegrip en tekstbegriptoetsen*. Arnhem: Dutch National Institute for Educational Measurement (CITO). Algemene publicatie, nr. 37.
- Vermunt, J.D.H.M. (1992). *Leerstijlen en sturen van leerprocessen in het hoger onderwijs*. Doctoral dissertation Tilburg University (KUB).
- VOG-MVT (1995). *Advies examenprogramma's havo/vwo. Moderne vreemde talen*. Enschede: SLO.
- Westhoff, G.J. (1981). *Voorspellend lezen, een didaktische benadering van de leesvaardigheidstraining in het moderne-vreemdetalenonderwijs*. Doctoral dissertation Utrecht University (RUU). Groningen: Wolters-Noordhoff.
- Westhoff, G.J. (1989). Some common teaching practices and their effectiveness in foreign language reading instruction. *European Journal of Teacher Education*, 12, (2), 121-130.
- Westhoff, G.J. (1993). Onderwijs in leesstrategieën: Theoretische achtergrond, praktisch nut. *Levende Talen*, 480, 266-270.
- Widdowson, H. (1979). The process and the purpose of reading. In H. Widdowson (Ed.), *Explorations in applied linguistics* (pp. 171-183). New York: Oxford University Press.
- Wildman, D.M. & Kling, M. (1979). Semantic, syntactic and spatial anticipation in reading. *Reading Research Quarterly*, 2, 128-164.
- Williams, R. (1983). Teaching the recognition of cohesive ties in reading a foreign language. *Reading in a Foreign Language*, 1 (1), 35-53.
- Wolf, D. (1993). *The relationship between word class and word meaning inferencing strategies in foreign language reading*. Paper presented at AILA 93, Amsterdam.
- Yorio, C.A. (1971). Some sources of reading problems for foreign language learners. *Language Learning*, 21, 107-115.
- Zuylén, N. van (1991). *Meerkeuze-clozetoetsen voor tekstbegrip Engels op C- en D-niveau. Verslag van een proefafname*. Arnhem: Dutch National Institute for Educational Measurement (CITO).

List of Appendices

Appendix I.	Pilot experiment: Research design	137
Appendix II.	Pilot experiment: Statistics	138
Appendix III.	Pilot experiment: Pre-test results	144
Appendix IV.	Main experiment: Subdivision into school classes	145
Appendix V.	Main experiment: Division of subjects over classes and conditions	146
Appendix VI.	Main experiment: Planning	147
Appendix VII.	Contents of the Spanish and English training programmes	148
Appendix VIII.	Two texts of the training programmes	150
Appendix IX.	Contents of the Spanish and English control programmes	157
Appendix X.	Two texts of the control programmes	159
Appendix XI.	Sample of tests	166
Appendix XII.	Statistics tryout vocabulary tests	178
Appendix XIII.	Questionnaires: Attitude, Self-confidence, Context use and Reading Ability	179
Appendix XIV.	Results by school class	191
Appendix XV.	Results reader-dependent variables	195
Appendix XVI.	Questionnaire: Evaluation of Programmes and Tests	198
Appendix XVII.	Results questionnaire Evaluation of Programmes and Tests	202
Appendix XVIII.	Main experiment: Statistics	214

Appendix I. Pilot experiment: Research design

	Pre-tests				Post-tests	
	S	E	S	E	S	E
class 1	O	O	T	C	O	O
class 2	O	O	C	T	O	O
class 3	O	O	C	C	O	O
class 4	O		C		O	
class 5		O		C		O
class 6		O		C		O
class 7		O		C		O

S = Spanish

E = English

T = Training Programme in Context-use

C = Control Reading Programme

O = Observation (pre- or post-tests)

Appendix II. Pilot experiment: Statistics**A. SPANISH****Spanish - Context tests****Table 1. Spanish: Context pre-test**

number of students	69		
number of items	20		
	mean	s.d.	
raw score	10.57	5.72	
item difficulty	.53		
reliability (Cronbach's alpha)	.73		

Table 2. Spanish: Context post-test

number of students	68		
number of items	20		
	mean	s.d.	
raw score	10.83	3.78	
item difficulty	.54		
reliability (Cronbach's alpha)	.78		

Spanish - Reading comprehension tests**Table 3. Spanish: Multiple-choice cloze pre-test**

number of students	87			
number of items	25			
chance score	8.33			
	mean	s.d.	lowest	highest
raw score	18.38	3.70	8.00	25.00
item difficulty	.74	.20	.26	.95
item-rest correlation	.29	.09	.14	.53
reliability (Cronbach's alpha)	.75			

Table 4. Spanish: Texts with multiple-choice questions pre-test

number of students	84			
number of items	25			
chance score	7.67			
	mean	s.d.	lowest	highest
raw score	20.75	3.15	8.00	25.00
item difficulty	.83	.10	.63	.99
item-rest correlation	.25	.11	.06	.47
reliability (Cronbach's alpha)	.70			

Table 5. Spanish: Multiple-choice cloze post-test

number of students	85			
number of items	25			
chance score	8.33			
	mean	s.d.	lowest	highest
raw score	15.34	5.02	5.00	25.00
item difficulty	.61	.11	.41	.82
item-rest correlation	.34	.16	-.03	.54
reliability (Cronbach's alpha)	.81			

Table 6. Spanish: Texts with multiple-choice questions post-test

number of students	43			
number of items	18			
chance score	5.67			
	mean	s.d.	lowest	highest
raw score	12.49	3.02	4.00	18.00
item difficulty	.69	.17	.16	.93
item-rest correlation	.27	.12	-.13	.41
reliability coefficient (Cronbach's alpha)	.68			

B. ENGLISH**English - Context tests****Table 7. English: Context pre-test**

number of students	136		
number of items	20		
	mean	s.d.	
raw score	11.27	3.83	
item difficulty	.56		
reliability coefficient (Cronbach's alpha)	.66		

Table 8. English: Context post-test

number of students	131		
number of items	20		
	mean	s.d.	
raw score	10.58	3.65	
item difficulty	.53		
reliability coefficient (Cronbach's alpha)	.66		

English - Reading comprehension tests**Table 9. English: Multiple-choice cloze pre-test**

number of students	152			
number of items	26			
chance score	8.67			
	mean	s.d.	lowest	highest
raw score	14.20	4.32	5.00	24.00
item difficulty	.55	.18	.14	.84
item-rest correlation	.26	.12	.03	.49
reliability (Cronbach's alpha)	.73			

Table 10. English: Texts with multiple-choice questions pre-test

number of students	160			
number of items	23			
chance score	6.45			
	mean	s.d.	lowest	highest
raw score	12.27	4.24	3.00	21.00
item difficulty	.53	.14	.25	.76
item-rest correlation	.28	.12	.05	.45
reliability (Cronbach's alpha)	.74			

Table 11. English: Multiple-choice cloze post-test

number of students	93			
number of items	25			
chance score	8.33			
	mean	s.d.	lowest	highest
raw score	17.40	3.13	9.00	23.00
item difficulty	.70	.19	.04	.99
item-rest correlation	.16	.14	-.16	.38
reliability (Cronbach's alpha)	.57			

Table 12. English: Texts with multiple-choice questions post-test

number of students	72			
number of items	21			
chance score	5.58			
	mean	s.d.	lowest	highest
raw score	14.43	2.69	9.00	20.00
item difficulty	.69	.19	.35	.99
item-rest correlation	.15	.18	-.24	.39
reliability coefficient (Cronbach's alpha)	.50			

Appendix III. Pilot experiment: Pre-test results

		School 1		School 2		
Spanish	n	Subj.		n	Subj.	
		mean	s.d		mean	s.d.
context test (items n =20)	33	9.83	5.66	36	11.25	5.77
reading test 1 (items n =25)	33	17.88	4.49	54	18.76	3.06
reading test 2 ** (items n =25)	31	19.32	3.55	53	21.47	2.76
English						
context test (items n =20)	58	14.73	2.98	78	14.71	4.11
reading test 1 (items n =26)	58	15.43	4.27	94	14.08	4.49
reading test 2 (items n =23)	61	12.49	4.38	99	12.65	3.59

**** = difference between scores of school 1 and school 2 significant at the 1% level**

Appendix IV. Main experiment: Subdivision into school classes

		Pre-tests				Post-tests	
		S	E	S	E	S	E
<hr/>							
Class:							
School 1							
Group 1	TS1	O		T		O	
	TS2	O		T		O	
Group 2	TE1		O		T		O
	TE2		O		T		O
	TE3		O		T		O
<hr/>							
School 2							
Group 3	CS1	O		C		O	
	CS2	O		C		O	
Group 4	CE1		O		C		O
	CE2		O		C		O

*S = Spanish**E = English**T = Training Programme in Context-use**C = Control Reading Programme**O = Observation (pre- or post-tests)*

Appendix V. Main experiment: Division of subjects over classes and conditions

		CLASS	Male	Female	Total
School 1	<i>Training</i>				
Spanish	TS1		10	8	18
	TS2		6	9	15
English	TE1		9	5	14
	TE2		4	10	14
	TE3		6	7	13
School 2	<i>Control</i>				
Spanish	CS1		11	11	22
	CS2		8	13	21
English	CE1		8	10	18
	CE2		12	6	18
		Total	74	79	153

Appendix VI. Main experiment: Planning

	1991	
May & June		Preparatory programme
July & August		Summer holidays
Pre-tests:		
9 - 13 September		first session pre-tests: -vocabulary test -context test
16 - 20 September		second session pre-tests: -reading comprehension test
Training and Control Programmes:		
23 September - 4 October		Lesson 1
7 - 11 October		October holidays
14 - 18 October		Lesson 2
21 - 25 October		Lesson 3
28 October - 8 November		Lesson 4
11 - 15 November		Lesson 5
18 - 22 November		Lesson 6
25 - 29 November		Lesson 7
		-questionnaire <i>Attitude</i>
Post-tests:		
2 - 6 December		first session post-tests: -vocabulary test -context test -questionnaire <i>Self-confidence</i> -questionnaire <i>Context use</i>
9 - 13 December		second session post-tests: -reading comprehension test -questionnaire <i>Reading comprehension</i>
16 - 20 December		spare week

Appendix VII. Contents of the Spanish and English training programmes

texts (number of words)	exercises	procedure	planned time
1 <i>El Metro de Madrid</i> (394)	<u>1.1</u> Mobilize prior knowledge	couples class	10 min.
<i>Hop on a bus, slowly</i>	<u>1.2</u> Text comprehension: match 6 given subtitles with the right text fragment	individual class	10
(401)	<u>1.3</u> Context use: (with help from rules) infer the meaning of 3 underlined words	class individual couples class	20
			40 min.
2 <i>El turismo, la primera industria de España</i> (331)	<u>2.1</u> Mobilize prior knowledge	class	10 min.
	<u>2.2</u> Text comprehension: matching exercise	individual class	10
	<u>2.3</u> Context use: logical connectors (survey)	class	10
<i>Swiss strike back with a big tick</i> (387)	<u>2.4</u> Context use: choose the correct logical connectors	individual class	10
			40 min.
3 <i>Del campo a la ciudad</i> (375)	<u>3.1</u> Mobilize prior knowledge	couples class	5 min.
	<u>3.2</u> Text comprehension: 6 true/false questions	individual class	10
	<u>3.3</u> Context use: logical connectors (survey)	class	10
	<u>3.4</u> Context use: choose the correct logical connector	individual couples class	5
<i>Future focus 1992</i> (389)	<u>3.5</u> Context use: infer the meaning of 3 underlined words	couples class	10
			40 min.

texts (number of words)	exercises	procedure	planned time
4 <i>La policía en el País Vasco</i> (410)	<u>4.1</u> Mobilize prior knowledge	individual	10 min.
<i>To find a job</i> (417)	<u>4.2</u> Context use: infer the meaning of 7 underlined words	individual couples class	30
			40 min.
5 <i>Amnistía Internacional</i> (372)	<u>5.1</u> Mobilize prior knowledge	individual	6 min.
	<u>5.2</u> Text comprehension: invent subtitles for the text fragments	individual class	10
<i>Old-fashioned law</i> (386)	<u>5.3</u> Context use: infer the meaning of 3 underlined words	individual class	12
	<u>5.4</u> Context use: underline the logical connectors and establish their meaning and function	individual class	12
			40 min.
6 <i>Barcelona</i> (302)	<u>6.1</u> Mobilize prior knowledge	individual couples	5 min.
	<u>6.2</u> Text comprehension: 10 true/false questions	individual class	10
<i>Worldwatch 2000</i> (376)	<u>6.3</u> Context use: infer the meaning of 5 underlined words	individual class	25
			40 min.
7 <i>Andalucía</i> (421)	<u>7.1</u> Mobilize prior knowledge	individual	5 min.
	<u>7.2</u> Context use: underline the logical connectors and explain their meaning and function	individual class	15
<i>Women's James Bond</i> (411)	<u>7.3</u> Context use: infer the meaning of 5 underlined words	individual class	15
			35 min.

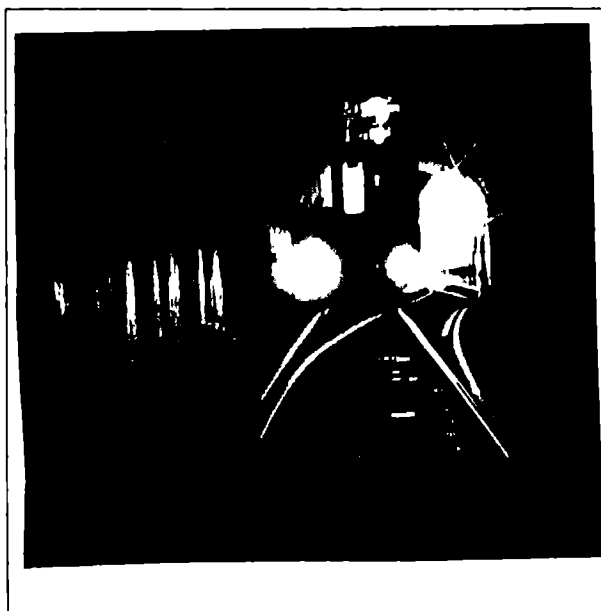
Appendix VIII. Two texts of the training programmes

(Spanish: lesson 1)

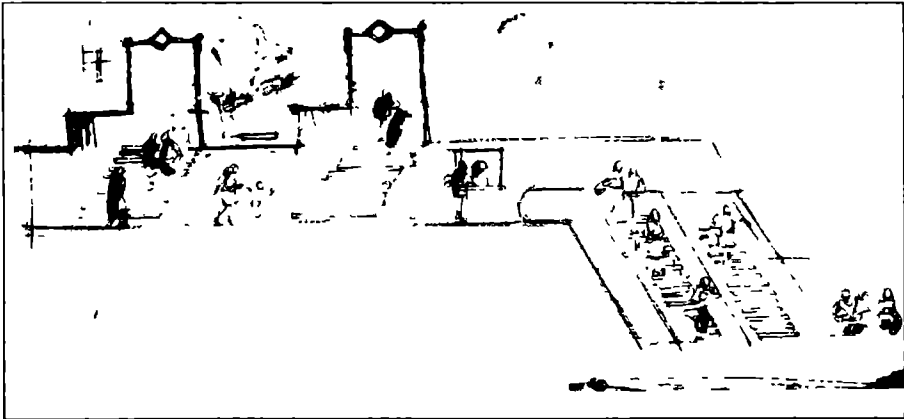
EL METRO DE MADRID

(I)

La capital de España, Madrid, tiene más de tres millones de habitantes. Todos estos habitantes y también las personas que visitan la ciudad, tienen que viajar por Madrid en coche, en autobús o... en Metro.

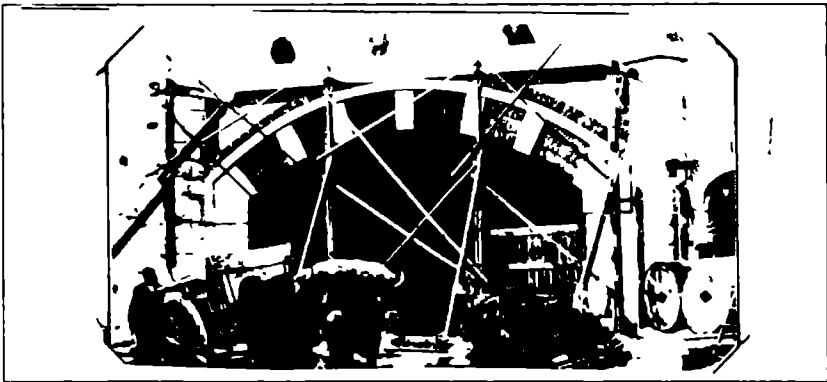


Por los túneles del Metro madrileño pasan todos los días muchos trenes para llevar a los viajeros de una estación a otra, de un punto de la ciudad a otro.



(II)

En 1917 se excava la primera línea del Metro: es un túnel de cuatro kilómetros que va de la estación del Sol a Cuatro Caminos. Después se han construido nuevas líneas, nuevas estaciones, nuevos trenes para unir las zonas más importantes de la ciudad.



(III)

El Metro de Madrid ha crecido mucho después de excavar la estación del Sol: hoy tiene una red de más de 100 kilómetros. Hay 10 líneas y más de 100 estaciones. Unas 5.000 personas trabajan en el Metro para los 300.000.000 de viajeros-¡casi 100 veces el número de habitantes de Madrid!

(IV)

Los trenes del Metro tienen un motor robusto, fuerte y alimentado por electricidad. La electricidad es una energía económica: por eso los trenes no cuestan mucho. Además no hay contaminación del aire en los túneles del Metro porque los motores eléctricos de los trenes son limpios: no sueltan humo como lo hacen los coches o los autobuses.

Cada tren tiene un conductor que lo conduce, controla la velocidad, abre y cierra las puertas paralos viajeros.

(V)

Todos los días los trenes del Metro funcionan más de 18 horas. Al terminar su trabajo a la una de la noche, los trenes van a una estación de servicio donde muchos técnicos los revisan.

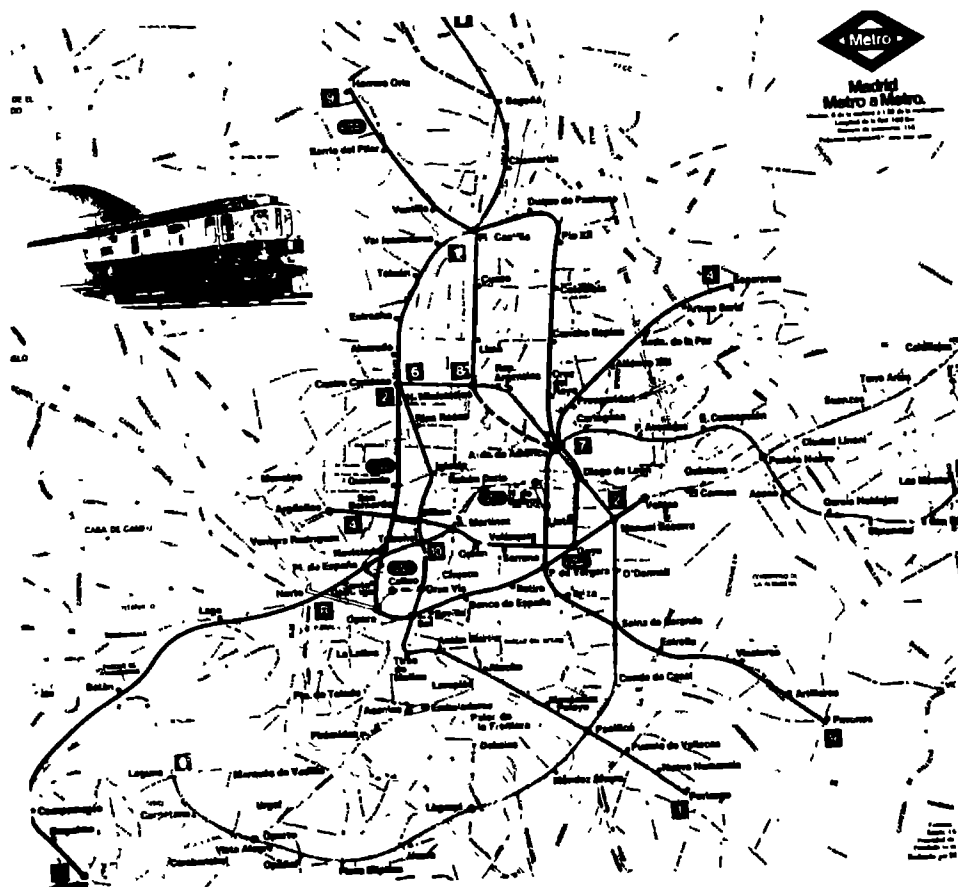
Entretanto otras muchas personas trabajan para limpiar los trenes y las estaciones del Metro.

Luego, cuando empieza el nuevo día, los trenes pueden transportar de nuevo a los viajeros a sus destinos.



(VI)

La historia del Metro no termina aquí: el Metro tiene un ayer, un hoy y un mañana. Pero ¿cómo es el futuro del Metro? No lo podemos saber. Lo que sí sabemos es que el Metro madrileño es un medio de transporte importante para el futuro de esta ciudad.



(English: lesson 2)

SWISS STRIKE BACK WITH A BIG TICK

The "Swatch": a big success in Great Britain

A Swiss plastic watch with a very large tick is challenging Asia's cheap market supremacy. The revolutionary quartz analogue "Swatch" (the name comes from Swiss watch), launched in Great Britain in 1983, is selling at a rate undreamed of by a Swiss watch industry noted more for its success at the luxury end of the timepiece business. Within a year, the initially budgeted

sales figure for Britain moved up from 50,000 to 200,000. ¹ Swiss production is still having trouble keeping up with English demand.



¹ And
But
For

The Swiss chose Britain as the first European market outside Switzerland, ____²____ success there means the watch will sell everywhere else.

Aggressive marketing: winning back markets lost to Japan

Swiss watchmakers readily admit that up till now the Japanese have beaten them on the marketing level. ____³____ with the Swatch they are hitting back with an aggressive campaign directed at the active, young consumer who wants to change watches with the latest fashion.

However, the Swatch's success is causing some heads to shake. Said a Swiss central banker who obliged his country's troubled watch industry by wearing it during the test period: 'I just couldn't stand that impossible tick. There was just no getting away from it'.

So far the Swatch has been introduced only in a few countries. ____⁴____, the Swiss have no doubt that their goal of over one million Swatch sales will be reached this year.

An offensive strategy

The big Swiss banks, led by Union Bank of Switzerland and Swiss bank Corporation, have temporarily taken over the major share of their country's watch industry. To avoid the bankruptcy of Switzerland's two largest watch concerns, ASUAG and SSIH, they arranged the merger of the two companies. The new ASUAG/SSIH group, of which the Swatch producer Eta SA is a part, is presently going through a huge production, sales and marketing reorganization to win back

² although
 because
 so

³ And
 But
 For

⁴ Besides
 Nevertheless
 Therefore

markets lost to Japan and Hong Kong over the last few years. ____⁵____ Switzerland invented the electronic watch, it was Japan which realized its potential and used the technology to challenge the European countries' position as the world's number one watchmaker. Now the Swiss watch industry is moving from a defensive to an offensive strategy.

Adapted from "*Financial Guardian*"

Timeless stars



5

Although
Because
Before

Appendix IX. Contents of the Spanish and English control programmes

texts (number of words)	exercises	procedure	planned time
1 <i>El Metro de Madrid</i> (394)	<u>1.1</u> Reading aloud: reading the text aloud (students); explaining unknown words (teacher)	class	15 min.
<i>Hop on a bus, slowly</i> (401)	<u>1.2</u> Text comprehension: 5 questions about text contents <u>1.3</u> Translation: the translation of two text paragraphs (orally)	class	10 15
			----- 40 min.
2 <i>El turismo, la primera industria de España</i> (331)	<u>2.1</u> Reading aloud: first paragraph (teacher); rest of text (students); explanation of unknown by the teacher <u>2.2</u> Text comprehension: matching exercise <u>2.3</u> Text comprehension: 6 questions about text contents	class	15 min.
<i>Swiss strike back with a big tick</i> (378)			----- 40 min.
3 <i>Del campo a la ciudad</i> (375)	<u>3.1</u> Reading aloud: reading the text aloud (students); explaining unknown words (teacher) <u>3.2</u> Text comprehension: 6 true/false questions <u>3.3</u> Translation: the translation of two text paragraphs (orally)	class	15 min.
<i>Future focus 1992</i> (389)			----- 40 min.

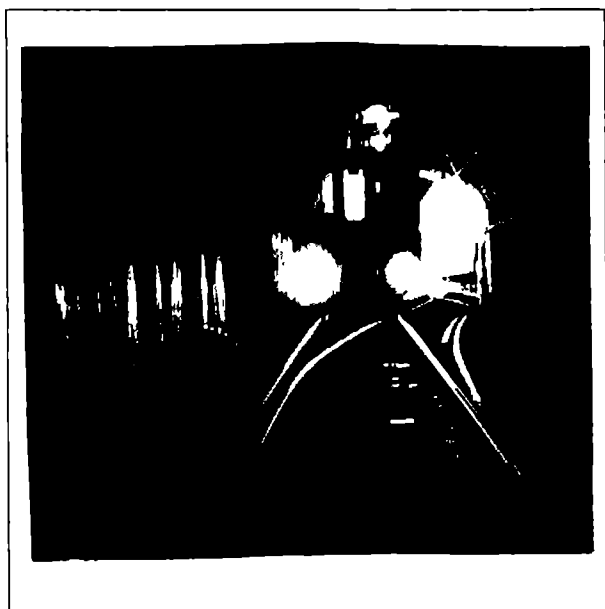
texts (number of words)	exercises	procedure	planned time
4 <i>La policía en el País Vasco</i> (410) <i>To find a job</i> (417)	<u>4.1</u> Reading aloud: reading the text aloud (students); explanation of unknown words (teacher) <u>4.2</u> Text comprehension: 7 questions about text contents	class individual class	20 min. 20 40 min.
5 <i>Amnistía Internacional</i> (372) <i>Old-fashioned law</i> (386)	<u>5.1</u> Reading aloud: reading the text aloud (students); explaining unknown words (teacher) <u>5.2</u> Translation: the translation of the (parts of) the text (orally) <u>5.3</u> Summary: retelling the main points of the text in own words	class class individual class	15 min. 15 10 40 min.
6 <i>Barcelona</i> (302) <i>Worldwatch 2000</i> (376)	<u>6.1</u> Reading aloud: reading the text aloud (students); explaining unknown words (teacher) <u>6.2</u> Text comprehension 10 true /false questions <u>6.3</u> Summary: retelling the main points of the text in own words	class class individual class	15 min. 10 15 40 min.
7 <i>Andalucía</i> (421) <i>Women's James Bond</i> (411)	<u>7.1</u> Reading aloud: reading the text aloud (students); explaining unknown words (teacher) <u>7.2</u> Translation: translation of the text	class class	15 min. 20 35 min.

Appendix X. Two texts of the control programmes

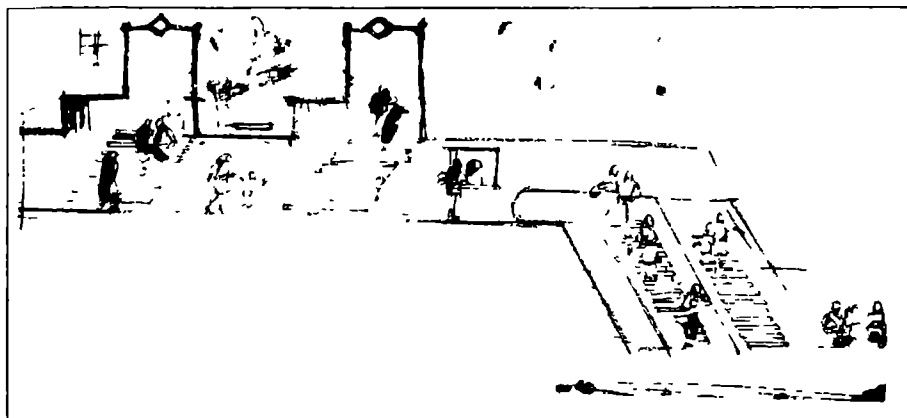
(Spanish: lesson 1)

EL METRO DE MADRID***El Metro, un medio de transporte***

La capital de España, Madrid, tiene más de tres millones de habitantes. Todos estos habitantes y también las personas que visitan la ciudad, tienen que viajar por Madrid en coche, en autobús o... en Metro.



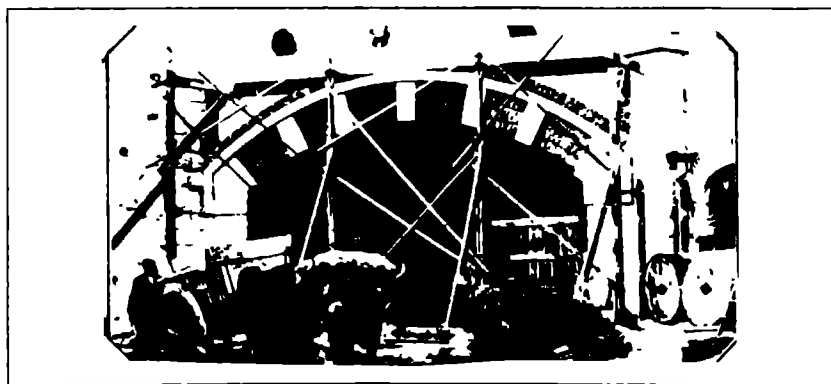
Por los túneles del Metro madrileño pasan todos los días muchos trenes para llevar a los viajeros de una estación a otra, de un punto de la ciudad a otro.



La historia del Metro madrileño

En 1917 se excava la primera línea del Metro: es un túnel de cuatro kilómetros que va de la estación del Sol a Cuatro Caminos. Después se han construido nuevas líneas, nuevas estaciones, nuevos trenes para unir las zonas más importantes de la ciudad.

El Metro, hoy: unas cifras



El Metro de Madrid ha crecido mucho después de excavar la estación del Sol: hoy tiene una red de más de 100 kilómetros. Hay 10 líneas y más de 100 estaciones. Unas 5.000 personas trabajan en el Metro para los 300.000.000 de viajeros-¡casi 100 veces el número de habitantes de Madrid!

¿Cómo funciona el Metro?

Los trenes del Metro tienen un motor robusto, fuerte y alimentado por electricidad. La electricidad es una energía económica: por eso los trenes no cuestan mucho. Además no hay contaminación del aire en los túneles del Metro porque los motores eléctricos de los trenes son limpios: no sueltan humo como lo hacen los coches o los autobuses.

Cada tren tiene un conductor que lo conduce, controla la velocidad, abre y cierra las puertas para los viajeros.

Las horas de trabajo del Metro

Todos los días los trenes del Metro funcionan más de 18 horas. Al terminar su trabajo a la una de la noche, los trenes van a una estación de servicio donde muchos técnicos los revisan.

Entretanto otras muchas personas trabajan para limpiar los trenes y las estaciones del Metro.

Luego, cuando empieza el nuevo día, los trenes pueden transportar de nuevo a los viajeros a sus destinos.



(English: lesson 2)

SWISS STRIKE BACK WITH A BIG TICK

The "Swatch": a big success in Great Britain

A Swiss plastic watch with a very large tick is challenging Asia's cheap market supremacy. The revolutionary quartz analogue "Swatch" (the name comes from Swiss watch), launched in Great Britain in 1983, is selling at a rate undreamed of by a Swiss watch industry noted more for its success at the luxury end of the timepiece business. Within a year, the initially budgeted sales figure for Britain moved up from 50,000 to 200,000. And Swiss production is still having trouble keeping up with English demand.



The Swiss chose Britain as the first European market outside Switzerland, because success there means the watch will sell everywhere else.

Aggressive marketing: winning back markets lost to Japan

Swiss watchmakers readily admit that up till now the Japanese have beaten them on the marketing level. But with the Swatch they are hitting back with an aggressive campaign directed at the active, young consumer who wants to change watches with the latest fashion.

However, the Swatch's success is causing some heads to shake. Said a Swiss central banker who obliged his country's troubled watch industry by wearing it during the test period: 'I just couldn't stand that impossible tick. There was just no getting away from it'.

So far the Swatch has been introduced only in a few countries. Nevertheless, the Swiss have no doubt that their goal of over one million Swatch sales will be reached this year.

An offensive strategy

The big Swiss banks, led by Union Bank of Switzerland and Swiss bank Corporation, have temporarily taken over the major share of their country's watch industry. To avoid the bankruptcy of Switzerland's two largest watch concerns, ASUAG and SSIH, they arranged the merger of the two companies. The new ASUAG/SSIH group, of which the Swatch producer Eta SA is a part, is presently going through a huge production, sales and marketing reorganization to win back

markets lost to Japan and Hong Kong over the last few years. Although Switzerland invented the electronic watch, it was Japan which realized its potential and used the technology to challenge the European countries' position as the world's number one watchmaker. Now the Swiss watch industry is moving from a defensive to an offensive strategy.

Adapted from "*Financial Guardian*"

Timeless stars



Appendix XI. Sample of tests

A. Context test (Spanish pre-test)

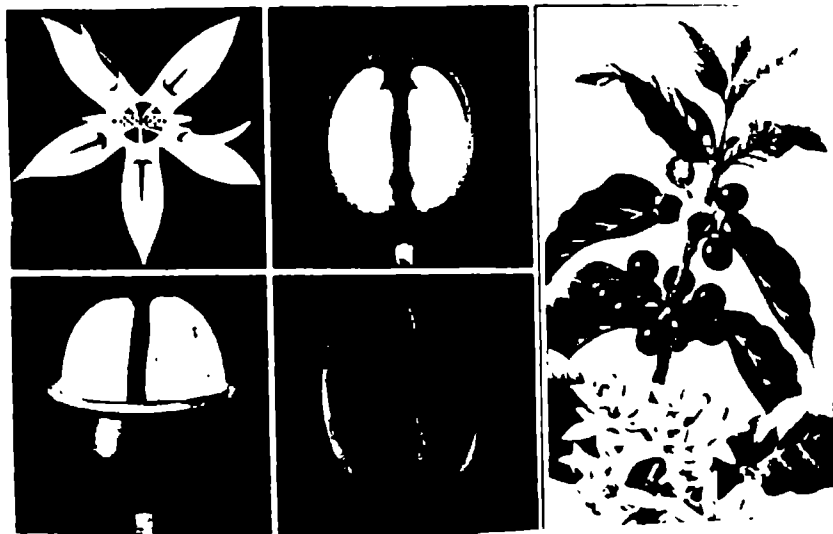
EL CAFÉ, BEBIDA INTERNACIONAL

El café producto importante para Latinoamérica

Uno de los productos más importantes para la economía de Latinoamérica es el café. En este continente el clima y el terreno son ideales para el cultivo de esta planta: hace mucho sol y terreno es relativamente **fértil**¹. Porque hay condiciones tan buenas, muchos países latinoamericanos como por ejemplo Brasil, Colombia, Ecuador y Guatemala **ocupan**² un puesto importante en la producción mundial del café.

La planta

La planta del café llega a uno o dos metros de **altura**³. Las **hojas**⁴ tienen un color verde oscuro pero las flores son blancas. La **madera** de la planta es dura y resistente y por eso la **usan**⁵ con frecuencia para fabricar **mesas**, sillas y otros muebles.



La producción del café

En general, **se cosecha**⁶ el café a mano. Después de cosechar el café lo ponen a **secar**⁷ al sol, lo limpian y después lo tuestan. La forma de tostar el café y el tipo de café determinan el **sabor**⁸ de la bebida.

Varios tipos de café

El café de Brasil, el primer producto mundial, es puro y concentrado y se cultiva en los terrenos bajos. Pero el café de Colombia es más suave⁹ y se cultiva en las laderas¹⁰ de las montañas. En Alemania combinamos el café brasileño con otros tipos de café más suave pero en Latinoamérica toman el café puro.

DESTRUCCIÓN Y CONSERVACIÓN DE LA NATURALEZA

La destrucción de la naturaleza por la caza¹ de animales

Los cocodrilos del Río Orinoco en Venezuela son los más grandes del mundo. En el siglo XIX, Alexander von Humboldt, especialista en animales y plantas de Sudamérica dice en sus libros que hay muchos cocodrilos en el Río Orinoco. Algunos de estos animales, así nos dice von Humboldt, tienen hasta siete metros de largo².

Pero nos dicen los zoólogos que ahora ya no hay muchos cocodrilos en los ríos de Venezuela. La piel del cocodrilo tiene mucho valor³ en el mercado mundial y por eso se matan millones de estos animales para producir bolsos y zapatos de su piel. En sólo dos años se exportan las pieles de 1.180.153 cocodrilos de Sudamérica, y muchos de ellas de menos de 60 centímetros de largo. Esta caza de animales para utilizar⁴ la piel no se limita a los cocodrilos. También otros animales que tienen una piel de valor, están en peligro⁵. Así hay ahora pocos jaguares y otros animales típicos de Sudamérica.



La destrucción de la naturaleza por la construcción de autopistas y ciudades

Otra causa de la destrucción de la naturaleza es la construcción de autopistas y ciudades en regiones donde hay muchas selvas⁶. Así desaparecen las grandes selvas del Amazonas. Porque estas selvas producen oxígeno, la destrucción de ellas va a tener efectos muy dañosos⁷ en el ambiente y en el clima. En la región amazónica va a disminuir la lluvia y a aumentar el calor y estos cambios de clima van a tener grandes consecuencias para el resto del continente americano y el mundo.



Los parques nacionales: un paso hacia la conservación de la naturaleza en Sudamérica

Hay muchos otros grandes problemas en Sudamérica, como por ejemplo la superpoblación de las ciudades, la ausencia⁸ de comida y las libertades inexistentes. Pero también el problema de la destrucción de la naturaleza preocupa⁹ a muchos suramericanos. Por eso varios países ahora tienen parques nacionales. Estos parques nacionales son grandes zonas en las que no está autorizada¹⁰ la construcción de autopistas, ciudades y aeropuertos. Gracias a estos parques se puede conservar un poco la rica y variada naturaleza suramericana.

Name:**Klasse:****Nummer:**

Raten Sie die unterstrichenen Wörter der zwei Texte.

El café, bebida internacional

	WORT	BEDEUTUNG
1.	fértil	_____
2.	ocupan	_____
3.	altura	_____
4.	hojas	_____
5.	usan	_____
6.	se cosecha	_____
7.	secar	_____
8.	sabor	_____
9.	suave	_____
10.	laderas	_____

Destrucción y conservación de la naturaleza

	WORT	BEDEUTUNG
1.	caza	_____
2.	de largo	_____
3.	valor	_____
4.	utilizar	_____
5.	peligro	_____
6.	selvas	_____
7.	dañinos	_____
8.	ausencia	_____
9.	preocupa	_____
10.	no está autorizada	_____

B. Reading comprehension test (English post-test)***The Missing Budgies Bureau***

When a budgie¹ dropped out of nowhere onto the shoulder of a woman strolling through a park, she was greatly surprised: it had blue and white feathers and a bright red face.

There was only one thing to do. She called on Alan Moon, the Sherlock Holmes of the budgie world. 'Can you _____(1)_____ its owner?', she asked.

Alan knew of a local factory where the girls in the first aid room kept a budgie. It had a habit of sitting on their shoulders and waiting for a kiss. That's why its face was usually covered in bright red _____(2)_____. Alan took the budgie to the factory and, _____(3)_____, it proved to be the girls' much-loved pet

It was all in a day's work for Alan, who has been running the Missing Budgies Bureau for the past 30 years. The chief aim of his labour of love is to return lost budgies to _____(4)_____

Alan, a 69-year-old retired power station worker, is a budgie fanatic _____(5)_____, he's been secretary of his local budgie society for more than a quarter of a century. In a back-garden cage he has 60 budgies of his own

Since he opened the bureau he has _____(6)_____ 2,300 cases of lost budgies. He has returned more than 800 to their owners. He has presented about 250 to _____(7)_____ after an unsuccessful search. The others have been found new homes.

Alan _____(8)_____ because of his detection system. If a budgie talks, Alan listens carefully to everything it says. If the budgie enjoys performing tricks - like walking upside-down - he makes a detailed note in his black register.

Recently a budgie flew through the window of a restaurant where guests were sitting down to a wedding reception. It landed on the table right in front of the bride and bridegroom. A waitress picked it up, popped it into a box and took it to _____(9)_____. The couple told Alan that if he couldn't trace the owner, they'd be _____(10)_____ to keep it themselves.

'Unfortunately,' says Alan, 'it was sick. I spent a week trying to cure it, but in the end I had to _____(11)_____. In the house I had another lost budgie whose owner I couldn't trace. I gave it to the bride and groom as a wedding present.'

In the world of Alan Moon, all budgie stories have a happy ending.

1 budgie = der Wellensittich

Travelling selfishly

Holidays in foreign parts, for the British, is no longer the privilege of the wealthy, and fully organized tours are more popular than ever before. But to take off, like Vanessa Brown did, for an indefinite period and go it alone is a different prospect altogether.

By many people's standards, Vanessa enjoyed a very successful life, but she was feeling increasingly dissatisfied. So she began to save and played with the idea of going away for a while. Her thoughts swam around the globe for a year until a friend broke a lunch date one day and she wandered into the Trailfinders shop in Earls Court Road to buy a book on India. An hour later she came out, somewhat dazed, one thousand pounds poorer and with tickets to fly her around the world in eight months. 'I got on a plane to Bombay in February. It was a strange sensation. I had no idea what I would do when I got there.'

Her reasons for going were quite clear. She wanted to know how she would react without the support of friends and work. Being freed of responsibilities can come as something of a shock and the first weeks of a long journey can be miserably disorientating. 'For the first two months I was really frightened and really homesick.' Being tall and fair, she hadn't a hope of remaining in the background in India and she experienced, for the first time, what it feels like to be racial minority.

Vanessa had no money problems. She allowed about £ 50 per week for accommodation, food and travel and, in the poorer countries, that was more than enough. She had left Britain with a large holdall of clothes, most of which she did not use, but had light, loose clothing made up as she went along. It is important to find a balance between an outfit that maintains your sense of well-being and one that enables you to see both the rough and the smooth of a place.

By the time Vanessa reached Kashmir, she was really beginning to relax. 'I had always been one of those people for whom life had never felt quite right. Now it did. It was a selfish, free sort of place. No telephones. No deadlines. The day was mine.'

Though it may sound contradictory. One of the most vital elements of any long trip is a holiday. Travelling is mentally and physically exhausting, and by the time Vanessa reached Delhi, at the end of May, in extreme heat, a holiday was exactly what she needed. Delhi was not the place to take it though. So, after a few weeks of trying to keep cool in the courtyard of a boarding house, she gave up and went to the beach ... in Thailand.

After visiting Hong Kong, Vanessa crossed the border into China, where her height, her fairness and round eyes made her into something of an oddity and the Chinese who lived in the country remained shyly in the background or came to touch her and to giggle. They were overwhelming hospitable and in one village a woman came out of a house and beckoned with a bowl of food. Taking up the invitation, she went inside and was amazed to see a television set broadcasting rugby from England! Moments like that reduce the world to sensible proportions and put one's life firmly back into perspective.

From an article by Phyllida Shaw in 'Company', January 1986

Scandal of a gap that lets children vanish

Everything seemed normal that Saturday morning as Lee Boxell set off for a typical teenager's day out - a trip to the shops and, later, a football match. At teatime, his family were waiting for the 15-year-old to push open the front door of their home in Cheam, Surrey. But he never returned.

That was 13 months ago. Lee is still missing, and the sorrow of two loving parents and a younger sister can only be imagined. The _____(23)_____ they face is experienced by a growing number of families. According to a report published tomorrow by the Children's Society, 98,000 children now go missing annually, an increase of 24,000 since the society's last survey two years ago.

The survey shows that many of the cases are _____(24)_____ reported to the police, in spite of the fact that some of those missing are as young as 13. Even when families do report the disappearance, the chances of finding the children are quite small. For Britain has no national register of missing people and files on the missing are kept only by individual police forces, with _____(25)_____ between the 52 forces nationwide.

Chief police officers want to improve the way missing people are _____(26)_____. In a report now with Douglas Hurd, the Home Secretary, the chief constables call for a computerized national missing persons register.

At present, the most advanced register is that held by the missing persons bureau at Scotland Yard. Roger Tideyman, who runs it, admits that the register is not fully exploited because other forces do not have to use it. 'If there are suspicious circumstances or if someone is particularly vulnerable, Scotland Yard is usually informed. _____(27)_____, the basic rule is that it is up to each police force to do as it sees fit, and so there is no real central control point.'

Last year the bureau held files on 9,951 missing people. By the end of the year, 2,156 were still missing. Of those reported missing, about 5,000 returned home of their own free will. Nearly a third of the above-mentioned _____(28)_____ were found as a result of police inquiries or publicity.

Many of the children who are found, or who give themselves up to the authorities, are not _____(29)_____ to return home. 16-year-old Kim, for instance, left her home in New Cumnock, Ayrshire, last May because she could not face the constant rows. She now lives in a comfortable 'safe house' in south London, opened by the Centrepont charity for the homeless specifically to house the growing number of 16-year-olds arriving _____(30)_____.

Kim is one of eight teenagers being looked after at the home, which aims to help them get on youth training schemes. Jacqueline Rea, the project leader, said there were many _____(31)_____ the children came to the home. 'It can be because of abuse and violence; it can be friction with a parent's girlfriend or boyfriend; it can be new children moving into the home.'

_____ (32)_____ Kim's mother knows that Kim is alive. The Boxells would dearly love to be in her shoes. They are now searching up and down the country for any clue to help them find their son. Christine Boxell still believes she may see Lee again. But she does not _____ (33)_____ that the police will find him. She said: 'When we called the police finally in the early hours of that Sunday, I thought Lee's face would be on television the same day and that the police would be out looking for him that night. But the police didn't really come until the next day. I was surprised that just because an older child was involved, the police didn't pay _____ (34)_____ attention to the case.'

'I would be happy just to know if Lee was alive, even if he didn't want to come back to us, but at the moment we just don't know anything at all,' she added _____ (35)_____.

The Missing Budgies Bureau

In dem Text wurden Wörter weggelassen. Wählen Sie das richtige weggelassene Wort (oder die weggelassenen Wörter) aus 3 Möglichkeiten.

Bitte nur eine Möglichkeit einsetzen.

- 1 A. find
B. phone
C. understood

- 10 A. forced
B. unable
C. very happy

- 2 A. blood
B. lipstick
C. paint

- 11 A. give it to the couple
B. put it to sleep
C. set it free

- 3 A. indeed
B. to Alan's surprise
C. unfortunately

Travelling selfishly

Bitte nur eine Möglichkeit einsetzen.

- 4 A. their finders
B. their owners
C. the Missing Budgie Bureau

12. **What is the main point in the first paragraph?**

- A. Going on holiday on one's own for an unfixed period is still rather unusual.
B. More and more people can afford to go on holiday in very far-away countries.
C. Only rich Britons can afford to go on holiday for a long time.
D. The British like going on holiday in groups more and more.

- 5 A. Luckily
B. Not surprisingly
C. Strangely

- 6 A. handled
B. heard of
C. solved

13. **What becomes clear about Vanessa from lines 5-9 ('By ... India')?**

- 7 A. the bureau
B. the finders
C. the owners

- A. She decided to give up her career although many people advised her against it.
B. She had a good income, but wanted to find out what living in primitive circumstances was like.
C. She was so disappointed in her friend that she wanted to be alone for some time

- 8 A. earns a lot of money
B. finds many new homes
C. is so successful

- D. She was thinking of making a long journey because she felt less and less happy about her situation.

- 9 A. Alan's home
B. its owner
C. the couple's table

14. **Why was Vanessa 'somewhat dazed' (line 9-10)?**
- A. She became aware that she would have to give up her situation.
 - B. She had booked an even longer journey than she had come for.
 - C. She had not realized there were so many books on India.
 - D. She had suddenly spent a lot of money by booking a long journey.
15. **What lesson is contained in lines 14-17 ('Her ... disorientating.')?**
- A. It may be unwise to go on a long tour without discussing it with your friends and colleagues.
 - B. People in general do not understand why someone should want to go on a holiday alone.
 - C. Single travellers may find it difficult to get used to their total independence.
 - D. The drawbacks of a long holiday only become clear after a few weeks.
16. **From lines 18-19 ('Being ... minority.') it becomes clear that**
- A. People in India do not really like tourists.
 - B. Vanessa was often treated very well in India because she was British.
 - C. Vanessa was very much aware in India of being a stranger.
17. **The word 'one' (line 21) refers to**
- A. 'balance' (line 24)
 - B. 'outfit' (line 24)
 - C. 'sense of well-being' (line 24)
18. **What advice is given in lines 23-25 ('It is ... a place.')?**
- A. You should also visit less well-known parts of the country in order to get the most out of your holidays.
 - B. You should always try to adapt yourself to the customs of the country you are visiting.
 - C. You should only take the clothes you will feel both comfortable and safe in.
 - D. You should only undertake those activities which you are capable of.
19. **What does Vanessa say about 'Kashmir' (line 26)?**
- A. It was a place where she could really do what she liked.
 - B. She wanted to stay there for the rest of her life.
 - C. She was treated quite well there.

20. **'Though it may sound contradictory' (line 29)
What may sound contradictory?**
- A. Doing nothing at all made Vanessa more tired than working.
- B. Travelling around relaxed Vanessa more than staying at one place.
- C. Vanessa did not like the hot weather while travelling.
- D. Vanessa needed time off while on holiday.
21. **How did the Chinese approach Vanessa, according to lines 35-39 ('After ... food')?**
- A. They often showed great friendliness.
- B. They pitied the sight of her and offered her food and shelter.
- C. They thought she looked ridiculous and laughed at her quite a few times.
- D. They were rather eager to invite a foreigner and tried hard to attract her attention.
22. **Which of the following expressions illustrates the contents of lines 39-42 ('Taking ... perspective.') best?**
- A. All's well that ends well.
- B. Home is where the heart is.
- C. It's a small world.

Scandal of a gap that lets children vanish

In dem folgenden Text wurden Wörter weggelassen. Wählen Sie das richtige weggelassene Wort (oder die richtigen weggelassenen Wörter) aus 3 Möglichkeiten.

Bitte nur eine Möglichkeit einsetzen.

- 23 A. danger
 B. shame
 C. unhappiness
- 24 A. immediately
 B. not even
 C. only
- 25 A. a lot of competition
 B. close co-operation
 C. little communication
- 26 A. counted
 B. helped
 C. traced
- 27 A. Also
 B. However
 C. Therefore
- 28 A. 2,156
 B. 5,000
 C. 9,951
- 29 A. afraid
 B. free
 C. willing
- 30 A. at the safe house
 B. back home
 C. in the capital city

31 A. inquiries when
 B. problems after
 C. reasons why

32 A. At least
 B. Not even
 C. Only

33 A. doubt
 B. expect
 C. hope

34 A. any
 B. immediate
 C. less

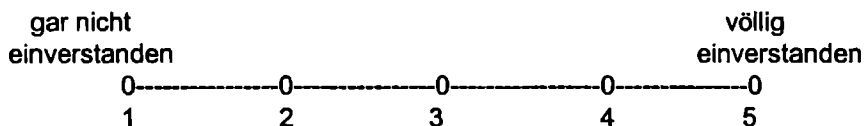
35 A. desperately
 B. heartlessly
 C. optimistically

Appendix XII. Statistics tryout vocabulary tests**Table 1. Spanish vocabulary test**

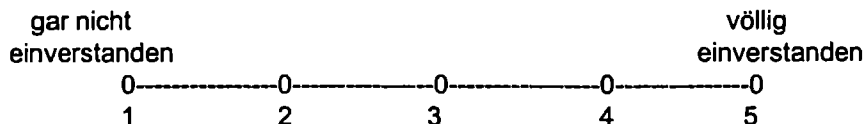
number of students	71			
number of items	50			
chance score	10.00			
	mean	s.d.	lowest	highest
raw score	42.52	5.00	27.00	49.00
item difficulty	.85	.14	.42	1.00
item-rest correlation	.24	.21	-.15	.60
reliability (Cronbach's alpha)	.80			

Table 2. English vocabulary test

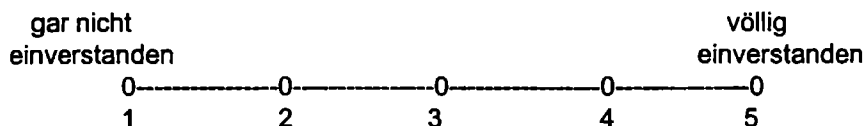
number of students	72			
number of items	50			
chance score	10.00			
	mean	s.d.	lowest	highest
raw score	29.50	7.43	12.00	46.00
item difficulty	.59	.24	.08	1.00
item-rest correlation	.30	.18	-.34	.62
reliability coefficient (Cronbach's alpha)	.85			



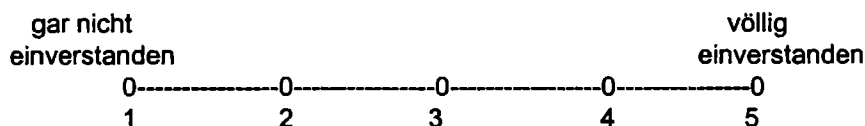
7 Es macht Spaß, Spanisch / Englisch zu lernen.



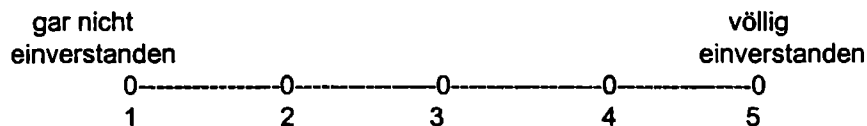
8 Es gibt in der Schule viele Sachen, die wichtiger zum lernen sind als Spanisch / Englisch.



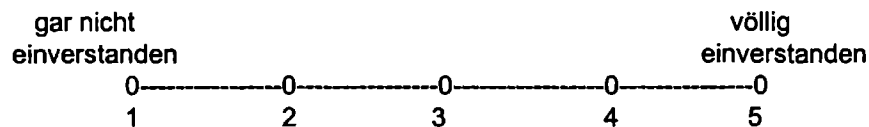
9 Eigentlich ist das Lernen von Spanisch / Englisch unangenehm.



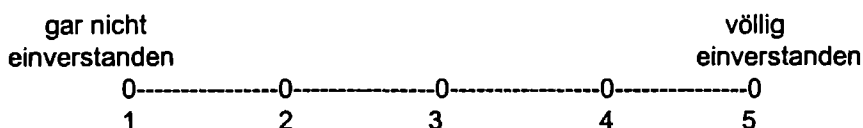
10 Ich wurde in der Schule Spanisch / Englisch lernen, auch wenn dies nicht obligatorisch wäre.



11 In der Schule hat das Lernen von Spanisch / Englisch keinen Zweck.



17 Ich komme meistens unvorbereitet in den Spanisch-Unterricht / Englisch-Unterricht.



2. Questionnaire Self-confidence

Fragebogen SV

Name:

Klasse:

Fach:

Nummer:

ANTWORTEN: Kreuzen Sie bitte die Antwort an, die auf Sie zutrifft. Wenn Sie keine passende Antwort finden, kreuzen Sie dann die Antwort an, die am meisten zutrifft.

Überschlagen Sie bitte keine Fragen.

Kreuzen Sie bitte nur eine Antwort an.

Es gibt bei diesem Fragebogen keine guten oder falschen Antworten.

SITUATION: Sie bekommen gleich zwei Lesetexte, in den Wörter unterstrichen sind. Diese Wörter haben Sie noch nicht gelernt. Sie kennen sie also nicht. Man wird Sie gleich bitten, die Bedeutung dieser Wörter dem Rest des Satzes oder des Textes zu entnehmen.

- 1 Wenn Sie mit so einer Aufgabe anfangen, kommt es dann vor, daß Sie das Gefühl haben, daß Sie die Aufgabe nicht schaffen können?

Das Gefühl, daß ich nicht weiß, ob ich es wohl kann, kommt bei mir

sehr oft oft manchmal fast nicht nie
vor vor vor vor vor
0-----0-----0-----0-----0

wenn ich mit so einer Aufgabe anfangen.

- 2 Wie gut meinen Sie, daß Sie solche Aufgaben machen?

Ich glaube, daß ich sie

sehr gut mittelmäßig nicht so sehr
gut gut schlecht
0-----0-----0-----0-----0

mache.

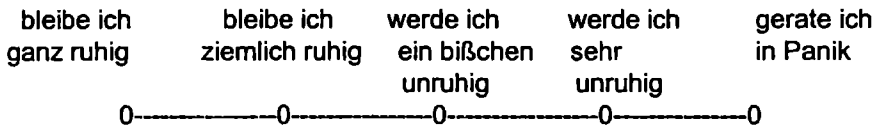
- 3 Einige Ihrer Kommilitonen zweifeln schnell, ob es Ihnen wohl gelingen wird, für eine solche Aufgabe eine ausreichende Note zu bekommen. Haben Sie das auch?

Zweifel, ob es mir wohl gelingen wird, für eine solche Aufgabe eine ausreichende Note zu bekommen, kommen mir

nie fast manchmal ziemlich sehr
 nie oft oft
0-----0-----0-----0-----0

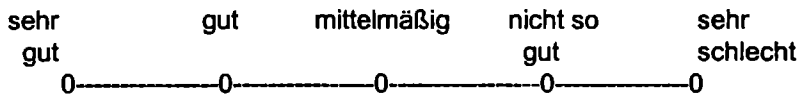
- 4 Einige Ihrer Kommilitonen geraten in Verwirrung, wenn sie beim Lesen auf unbekannten Wörter stoßen. Was würden Sie machen?

Wenn ich beim Lesen auf unbekannte Wörter stoße,



- 5 Es werden Fragen über einen Text gestellt. Wie gut meinen Sie, daß Sie solche Fragen beantworten können?

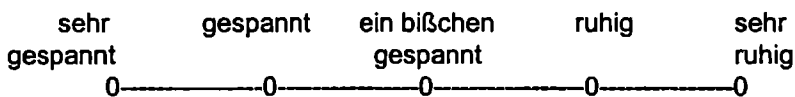
Ich denke, daß ich das



kann.

- 6 Ihr Lehrer / Ihre Lehrerin gibt Ihnen nächstes Mal Ihre Lesearbeit zurück. Wie meinen Sie, daß Sie sich beim Zurückgeben fühlen?

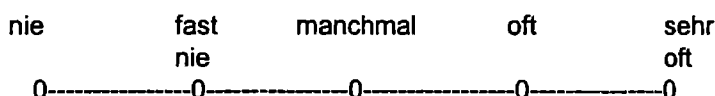
Ich glaube, daß ich mich beim Zurückgeben



fühlen werde.

- 7 Sie habe Ihre Arbeit zurück bekommen. Gesetz den Fall, das Resultat ist schlechter, als Sie erwartet hatten. Kommen Ihnen in so einer Situation Zweifel, ob Sie es wohl können?

In so einer Situation zweifle ich



daran, ob ich es kann.

3. Questionnaire Context use

Fragebogen KO

Name:

Klasse:

Fach:

Nummer:

ANTWORTEN: Kreuzen Sie bitte die Antwort an, die auf Sie zutrifft. Wenn Sie keine passende Antwort finden, kreuzen Sie dann die Antwort an, die am meisten zutrifft.

Überschlagen Sie bitte keine Fragen.

]

Kreuzen Sie bitte nur eine Antwort an.

Wir möchten wissen wie Sie bei der Lösung der Leseaufgabe, die Sie soeben gemacht haben, vorgegangen sind. Dafür ist es sehr wichtig, daß Sie genau angeben wie Sie vorgegangen sind und nicht wie Sie vorgehen wollten.

Es gibt bei diesem Fragebogen keine guten oder falschen Antworten.

Wenn ich glaubte die Bedeutung des unterstrichenen Wortes zu kennen,

- 6 überlegte ich, ob diese Bedeutung, was den Inhalt betraf in den Text paßte.

nie fast nie manchmal fast immer immer
 0-----0-----0-----0-----0

- 7 überlegte ich, ob diese Bedeutung, was die Funktion (z. B. Subjekt, Prädikat, Akkusativobjekt usw.) betraf, in den Satz paßte.

nie fast nie manchmal fast immer immer
 0-----0-----0-----0-----0

- 8 überlegte ich, ob diese Bedeutung, was die Wortart (z. B. Substantiv, Adjektiv, Verb oder Adverb) betraf, in den Satz paßte.

nie fast nie manchmal fast immer immer
 0-----0-----0-----0-----0

4. Questionnaire Reading Ability

Fragebogen LF

Name:

Klasse:

Fach:

Nummer:

ANTWORTEN: Kreuzen Sie bitte die Antwort an, die auf Sie zutrifft. Wenn Sie keine passende Antwort finden, kreuzen Sie dann die Antwort an, die am meisten zutrifft.

Überschlagen Sie bitte keine Fragen.

Kreuzen Sie bitte nur eine Antwort an.

Es gibt bei diesem Fragebogen keine guten oder falschen Antworten.

- 1 Wie gut finden Sie sich im Lesen eines Textes in dieser modernen Fremdsprache?

Ich finde, daß ich

gut	ziemlich gut	nicht so gut	ziemlich schlecht	sehr schlecht
0-----0-----0-----0-----0				

bin im Lesen eines Textes in dieser modernen Fremdsprache.

- 2 Wie schnell meinen Sie, daß Sie einen Text in dieser modernen Fremdsprache lesen?

Ich finde, daß ich einen Text

schnell	ziemlich schnell	nicht so schnell	ziemlich langsam	sehr langsam
0-----0-----0-----0-----0				

lese.

- 3 Glauben Sie, daß Sie die groben Umrissse eines Textes in dieser modernen Fremdsprache verstehen können?

Ich finde, daß ich die groben Umrissse eines Textes

immer	fast immer	manchmal	fast nie	nie
0-----0-----0-----0-----0				

verstehen kann.

- 4 In welchem Maße meinen Sie, daß Sie in einem Text in dieser modernen Fremdsprache Haupt- und Nebensachen unterscheiden können?

Ich finde, daß ich in einem Text

gut	ziemlich gut	nicht so gut	ziemlich schlecht	sehr schlecht
0-----0-----0-----0-----0				

Haupt- und Nebensachen unterscheiden kann.

- 5 In welchem Maße meinen Sie, daß Sie einen Text in dieser modernen Fremdsprache verstehen?

Ich finde, daß ich einen Text

schnell	ziemlich schnell	nicht so schnell	ziemlich langsam	sehr langsam
0-----0-----0-----0-----0				

verstehe.

Appendix XIV. Results by school class

Spanish - Context tests

Table 1. Context tests for Spanish: Means and standard deviations (proportions)

	Subj.	pre-test (items n= 20)		post-test (items n= 20)		gain scores	
		n	mean s.d.	mean s.d.		mean s.d.	
TS1	15	.31	.22	.40	.18	.09	.13
TS2	10	.48	.23	.46	.15	-.02	.14
CS1	10	.87	.07	.79	.10	-.08	.12
CS2	14	.72	.13	.66	.14	-.06	.13

TS1 = Spanish Training Programme - School 1, Class 1

TS2 = Spanish Training Programme - School 1, Class 2

CS1 = Spanish Control Programme - School 2, Class 1

CS2 = Spanish Control Programme - School 2, Class 2

Table 2. Context tests for Spanish: Results of the t-tests on the mean gain scores

School classes	t-value	Degrees of Freedom	2-tail probability
TS1 vs. CS1	3.37	23	.00**
TS1 vs. CS2	3.05	27	.01**
TS2 vs. CS1	1.13	18	.27
TS2 vs. CS2	.75	22	.46
TS1 vs. TS2	1.92	23	.07
CS1 vs. CS2	-.48	22	.64

**** = significant at the 1% level**

Spanish - Reading comprehension tests

Table 3. Reading comprehension tests for Spanish: Means and standard deviations (proportions)

	Subj.	pre-test (items n=34)		post-test (items n=35)		gain scores	
		n	mean s.d.	mean s.d.		mean s.d.	
TS1	15		.47 .15	.50 .12		.03 .10	
TS2	10		.64 .13	.28 .06		-.36 .18	
CS1	10		.65 .09	.71 .07		.06 .08	
CS2	14		.66 .10	.61 .12		-.05 .13	

TS1 = Spanish Training Programme - School 1, Class 1

TS2 = Spanish Training Programme - School 1, Class 2

CS1 = Spanish Control Programme - School 2, Class 1

CS2 = Spanish Control Programme - School 2, Class 2

Table 4. Reading comprehension tests for Spanish: Results of the t-tests on the mean gain scores

School classes	t-value	Degrees of Freedom	2-tail probability
TS1 vs. CS1	-.88	23	.39
TS1 vs. CS2	1.75	27	.09
TS2 vs. CS1	-6.83	12.70	.00**
TS2 vs. CS2	-5.07	22	.00**
TS1 vs. TS2	6.34	12.58	.00**
CS1 vs. CS2	2.30	22	.03*

**** = significant at the 1% level**

*** = significant at the 5% level**

English - Context tests**Table 5. Context tests for English: Means and standard deviations (proportions)**

	Subj.	pre-test (items n= 20)		post-test (items n= 20)		gain scores	
		n	mean s.d.	mean s.d.		mean s.d.	
TE1	12		.50 .20	.59 .11		.09 .17	
TE2	9		.58 .13	.46 .15		-.12 .09	
CE1	8		.60 .11	.54 .07		-.06 .10	
CE2	10		.62 .24	.51 .28		-.11 .17	

TE1 = English Training Programme - School 1, Class 1

TE2 = English Training Programme - School 1, Class 2

CE1 = English Control Programme - School 2, Class 1

CE2 = English Control Programme - School 2, Class 2

Table 6. Context tests for English: Results of the t-tests on the mean gain scores

School classes	t-value	Degrees of Freedom	2-tail probability
TE1 vs. CE1	2.27	18	.04*
TE1 vs. CE2	2.79	20	.01**
TE2 vs. CE1	-1.28	15	.22
TE2 vs. CE2	-.11	17	.91
TE1 vs. TE2	3.38	19	.00**
CE1 vs. CE2	.75	16	.46

**** = significant at the 1% level**

*** = significant at the 5% level**

English - Reading comprehension tests**Table 7. Reading comprehension tests for English: Means and standard deviations (proportions)**

	Subj.	pre-test (items n=37)		post-test (items n=35)		gain scores	
		n	mean s.d.	mean s.d.		mean s.d.	
TE1	12	.57	.11	.51	.08	-.06	.12
TE2	9	.68	.10	.61	.14	-.07	.12
CE1	8	.54	.12	.52	.12	-.02	.09
CE2	10	.58	.14	.48	.16	-.10	.13

TE1 = English Training Programme - School 1, Class 1

TE2 = English Training Programme - School 1, Class 2

CE1 = English Control Programme - School 2, Class 1

CE2 = English Control Programme - School 2, Class 2

Table 8. Reading comprehension tests for English: Results of the t-tests on the mean gain scores

School classes	t-value	Degrees of Freedom	2-tail probability
TE1 vs. CE1	-.88	18	.39
TE1 vs. CE2	.77	20	.45
TE2 vs. CE1	-1.01	15	.33
TE2 vs. CE2	.54	17	.60
TE1 vs. TE2	.18	19	.86
CE1 vs. CE2	1.15	16	.15

****** = significant at the 1% level

***** = significant at the 5% level

Appendix XV. Results reader-dependent variables**Spanish - Vocabulary****Table 1. Spanish. Correlations of the results on the vocabulary post-test, context post-test and reading comprehension post-test**

	vocabulary	context use	reading
vocabulary	1.00 (n=48)	.64** (n=48)	.35** (n=48)
context use		1.00 (n=49)	.62** (n=49)
reading			1.00 (n=49)

**** = significant at the 1% level**

English - Vocabulary**Table 2. English. Correlations of the results on the vocabulary post-test, context post-test and reading comprehension post-test**

	vocabulary	context use	reading
vocabulary	1.00 (n=43)	.62** (n=43)	.66** (n=43)
context use		1.00 (n=43)	.53** (n=43)
reading			1.00 (n=43)

**** = significant at the 1% level**

Spanish - Attitude, estimate of context use and self-confidence

Table 3. Spanish. Correlations of the scores on the variables 'attitude', 'estimate of context use' and 'self-confidence' and the results on the context and reading comprehension tests

	context test	reading test	attitude	estimate of context use	self-confidence
context test	1.00 (n=49)	.62** (n=49)	.51** (n=48)	.11 (n=49)	.28* (n=49)
reading test		1.00 (n=49)	.32* (n=48)	.27 (n=49)	.03 (n=49)
attitude			1.00 (n=48)	.11 (n=48)	.06 (n=48)
estimate of context use				1.00 (n=49)	.21 (n=49)
self-confidence					1.00 (n=49)

** = significant at the 1% level

* = significant at the 5% level

English - Attitude, estimate of context use and self-confidence**Table 4. English. Correlations of the scores on the variables 'attitude', 'estimate of context use' and 'self-confidence' and the results in the context and reading comprehension tests**

	context test	reading test	attitude	estimate of context use	self-confidence
context- test	1.00 (n=43)	.53** (n=43)	.00 (n=37)	.10 (n=42)	-.04 (n=42)
reading test		1.00 (n=43)	.29 (n=37)	.25 (n=42)	.17 (n=37)
attitude			1.00 (n=37)	.29 (n=37)	.53** (n=37)
estimate of context use				1.00 (n=42)	.24 (n=42)
self-confidence					1.00 (n=42)

**** = significant at the 1% level**

Appendix XVI. Questionnaire: Evaluation of Programmes and Tests**Ihre Meinung über das Leseprogramm**

Wir möchten Ihre Meinung über das Leseprogramm als Ganzes (= Texte + Aufgaben + Tests) wissen. Sie helfen uns dann, das Programm zu verbessern.

ANTWORTEN: Kreuzen Sie bitte die Antwort an, die auf Sie zutrifft. Wenn Sie keine passende Antwort finden, kreuzen Sie dann die Antwort an, die am meisten zutrifft.

Überschlagen Sie bitte keine Fragen.

Kreuzen Sie bitte nur eine Antwort an.

Es gibt bei diesem Fragebogen keine guten oder falschen Antworten.

1 = gar nicht einverstanden

2 = nicht einverstanden

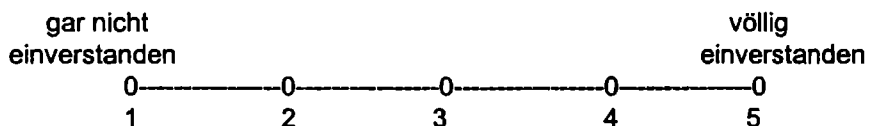
3 = ich zweifle zwischen einverstanden - nicht einverstanden

4 = einverstanden

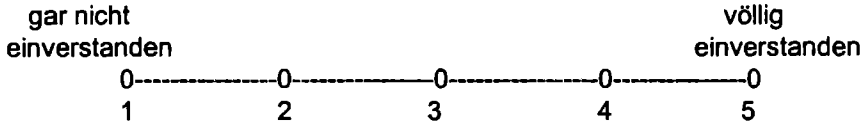
5 = völlig einverstanden

LESEPROGRAMM

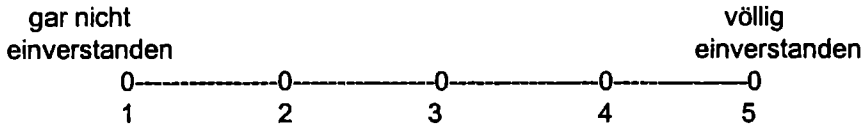
1 Ich fand das Programm nützlich, um zu lernen, wie ich die Bedeutung von Wörtern, die ich nicht kenne, dem Rest des Textes entnehmen kann.



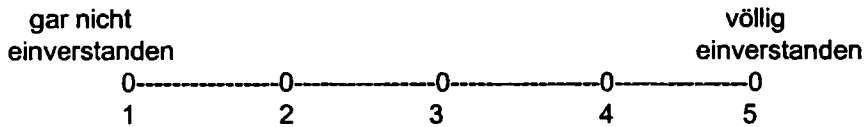
2 Ich fand es angenehm, am Leseprogramm zu arbeiten.



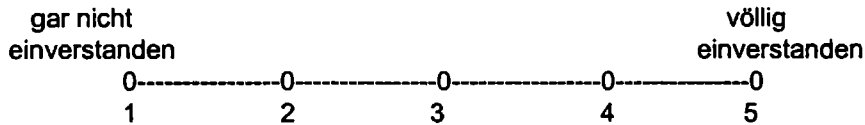
3 Ich finde, daß ich Texte in dieser Fremdsprache jetzt besser verstehen kann.



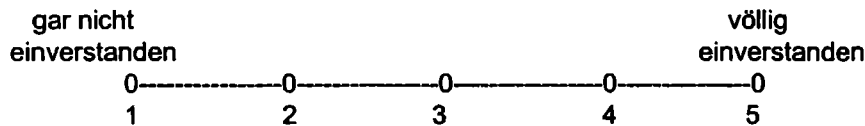
4 Die Absicht der Aufgaben des Programms war immer klar.



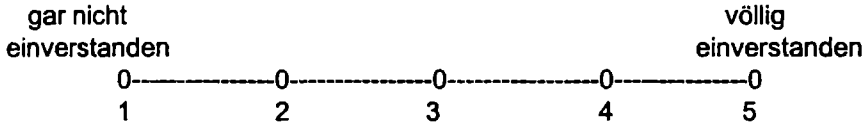
5 Ich fand, daß genügend Abwechslung im Programm war.



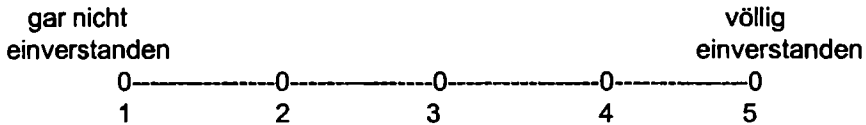
6 Die Ergebnisse der Aufgaben des Programms, waren für mich meistens enttäuschend.



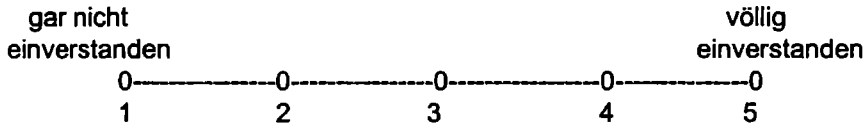
12 Ich hatte zu wenig Zeit, die Texte der Lesefertigkeits-Tests zu lesen.



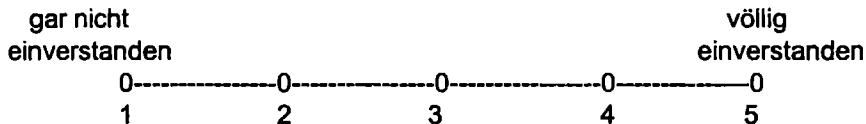
13 Ich hatte zu wenig Zeit, die Aufgaben der Wortschatz-Tests zu machen.



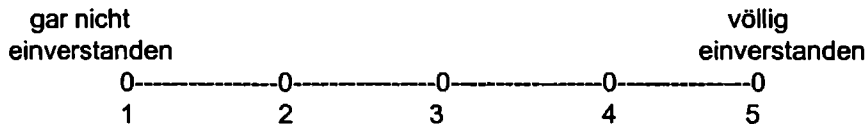
14 Ich hatte zu wenig Zeit, die Aufgaben der "Raten Sie das unterstrichene Wort" - Tests zu machen.



15 Ich hatte zu wenig Zeit, die Aufgaben der Lesefertigkeits-Tests zu machen.



16 Die Ergebnisse der Aufgaben der Tests, waren für mich meistens enttäuschend.



Appendix XVII. Results questionnaire Evaluation of Programmes and Tests

Table 1. Results by group. Mean scores and standard deviations

	Spanish				English			
	ST (n=30)		SC (n=31)		ET (n=33)		EC (n=34)	
<i>I. programmes</i>								
	\bar{x}	s.d	\bar{x}	s d	\bar{x}	s.d	\bar{x}	s d
1 useful programme	2.70	1.12	2.55	1.34	3.21	1.22	2.94	1.15
2 pleasant programme	2.27	1.08	2.29	1.27	2.51	1.09	2.79	1.17
3 improved comprehension	2.17	.98	2.32	1.04	2.88	1.08	2.47	.99
4 aim clear	3.07	1.17	2.87	1.50	3.06	1.17	3.23	1.07
5 enough diversity	2.77	1.13	2.61	1.36	2.58	1.27	3.15	1.18
6 results tasks	2.90	.92	2.84	1.27	3.24	1.27	3.00	1.35
7 time for reading texts	3.20	1.21	3.16	1.21	3.48	1.06	3.09	1.14
8 time for tasks	3.43	1.04	3.26	1.12	3.12	1.24	3.23	.89
9 teaching satisfying	2.80	.89	2.71	1.37	2.91	1.13	2.53	1.26
questions 1-9	2.81	1.06	2.73	1.27	3.00	1.17	2.94	1.13
<i>II. tests</i>								
10 aim clear	3.03	1.07	2.87	1.38	3.27	1.10	3.26	1.02
11 context time reading	3.10	.99	3.16	1.07	3.12	1.19	3.12	1.04
12 reading time reading	3.23	1.01	3.32	.91	3.54	.79	3.03	1.03
13 vocabulary. time	3.07	1.01	3.55	.96	3.67	.64	3.21	1.01
14 context time tasks	2.87	1.04	3.29	.94	3.48	.83	3.18	1.00
15 reading time tasks	3.03	.85	3.39	1.05	3.51	.71	3.09	.96
16 results tests	2.80	.85	3.19	1.14	3.27	1.04	2.94	1.15
questions 10-16	3.02	.97	3.25	1.06	3.41	.90	3.12	1.03
ST = Spanish Training Group - School 1						maximum value = 5		
SC = Spanish Control Group- School 2						minimum value = 1		
ET = English Training Group - School 1								
EC = English Control Group - School 2								

Table 2. Percentages of the frequencies of the negative, neutral and positive scores. Results by group

	Spanish ST (n=30)			SC (n=31)			English ET (n=33)			EC (n=34)		
<i>I. programmes</i>	-	±	+	-	±	+	-	±	+	-	±	+
1 useful programme	30 0	50 0	20 0	45 2	22 6	32 2	27 3	15 2	57 6	41 2	23 5	35 3
2 pleasant programme	63 3	23 3	13 3	54 8	19 4	25 8	45 4	33 3	21 2	44 1	32 4	23 6
3 improved comprehension	56 7	36 7	6 7	48 4	45 2	6 4	27 3	39 4	33 3	47 1	38 2	14 7
4 aim clear	30 0	26 7	43 4	35 5	16 1	48 4	36 4	21 2	42 4	29 4	17 6	53 0
5 enough diversity	40 0	30 0	30 0	45 2	22 6	32 3	48 5	24 2	27 3	29 4	32 4	38 2
6 results tasks	30 0	46 7	23 3	35 5	29 0	35 5	21 2	24 2	54 6	44 1	17 6	38 2
7 time for reading texts	26 7	33 3	40 0	25 8	22 6	51 7	21 2	21 2	57 6	26 5	35 3	38 2
8 time for tasks	16 7	36 7	46 7	19 4	38 7	41 9	30 3	30 3	39 4	14 7	44 1	41 1
9 teaching satisfying	30 0	60 0	10 0	45 2	16 1	38 8	30 3	39 4	30 3	50 0	20 6	29 4
<i>II. tests</i>												
10 aim clear	26 7	40 0	33 4	32 3	25 8	42 0	21 2	30 3	48 5	23 5	35 3	41 2
11 context time reading	16 7	53 3	30 0	22 6	35 5	42 0	30 3	24 2	45 5	26 5	38 2	35 3
12 reading time reading	13 3	46 7	40 0	16 1	38 7	45 2	9 1	36 4	54 6	32 4	32 4	35 3
13 vocabulary time	20 0	50 0	30 0	12 9	29 0	58 1	3 0	33 3	63 7	29 4	23 5	47 1
14 context time tasks	23 3	53 3	23 3	19 4	35 5	45 2	12 1	27 3	60 6	29 4	26 5	44 1
15 reading time tasks	13 3	60 0	26 7	16 1	35 5	48 4	6 1	42 4	51 6	23 5	44 1	32 4
16 results tests	26 7	66 7	6 7	19 4	48 4	32 2	21 2	33 3	45 5	41 2	23 5	35 3

ST = Spanish Training Group - School 1

SC = Spanish Control Group - School 2

ET = English Training Group - School 1

EC = English Control Group - School 2

- = negative score (1 and 2)

± = neutral score (3)

+

= positive score (4 and 5)

Table 3. Results by school class. Mean scores and standard deviations

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

Question:	<i>Entire Population</i>		<i>Spanish Training</i>		<i>Spanish Control</i>		<i>English Training</i>		<i>English Control</i>	
	\bar{x}	s.d	\bar{x}	s.d	\bar{x}	s.d	\bar{x}	s.d	\bar{x}	s.d
15. reading tests:										
time tasks	3.26	.92	3.03	.85	3.39	1.05	3.51	.71	3.09	.96
TS1			3.23	.75						
TS2			2.77	.93						
CS1					3.60	.83				
CS2					3.19	1.22				
TE1							3.33	.78		
TE2							3.56	.73		
TE3							3.67	.65		
CE1									3.23	.56
CE2									2.94	1.25
16. results tests	3.05	1.06	2.80	.85	3.19	1.14	3.27	1.04	2.94	1.15
TS1			2.88	.78						
TS2			2.69	.95						
CS1					3.33	.72				
CS2					3.06	1.44				
TE1							2.92	1.24		
TE2							3.44	1.01		
TE3							3.50	.80		
CE1									2.71	1.05
CE2									3.18	1.24

TS1 = Spanish Training Programme - School 1, Class 1

TS2 = Spanish Training Programme - School 1, Class 2

CS1 = Spanish Control Programme - School 2, Class 1

CS2 = Spanish Control Programme - School 2, Class 2

TE1 = English Training Programme - School 1, Class 1

TE2 = English Training Programme - School 1, Class 2

TE3 = English Training Programme - School 1, Class 3

CE1 = English Control Programme - School 2, Class 1

CE2 = English Control programme - School 2, Class 2

Table 4. Training condition. Percentages of the frequencies of the negative, neutral and positive scores. Results by school class

	Spanish TS1 (n=17)			TS2 (n=13)			English TE1 (n=12)			TE2 (n=9)			TE3 (n=12)		
<i>I. Programmes</i>	-	±	+	-	±	+	-	±	+	-	±	+	-	±	+
1 useful programme	29.4	52.9	17.6	30.8	46.2	23.1	75.0	16.7	8.3	0.0	22.2	77.8	0.0	8.3	91.6
2 pleasant programme	70.6	11.8	17.6	53.8	38.5	7.7	83.3	16.7	0.0	22.2	55.6	22.2	25.0	33.3	41.7
3 improved comprehension	58.8	35.3	5.9	53.8	38.5	7.7	50.0	41.7	8.3	22.2	44.4	33.3	8.3	33.3	58.3
4 purpose clear	17.6	23.5	58.9	46.2	30.8	23.1	58.3	25.0	16.7	33.3	33.3	33.3	16.7	8.3	75.0
5 enough diversity	47.1	17.6	35.3	30.8	46.2	23.1	83.3	8.3	8.3	44.4	33.3	22.2	16.7	33.3	50.0
6 results tasks	23.5	52.9	23.5	38.5	38.5	23.1	41.7	33.3	25.0	11.1	11.1	77.8	8.3	25.0	66.7
7 time for reading texts	23.5	11.8	64.7	30.8	61.5	7.7	33.3	25.0	41.6	22.2	11.1	66.6	8.3	25.0	66.6
8 time for tasks	17.6	29.4	52.9	15.4	46.2	38.5	50.0	41.7	8.3	11.1	11.1	77.8	25.0	33.3	41.7
9 teaching satisfying	17.6	70.6	11.8	46.2	46.2	7.7	50.0	33.3	16.7	22.2	55.6	22.2	16.7	33.3	50.0
<i>II. Tests</i>	-	±	+	-	±	+	-	±	+	-	±	+	-	±	+
10 purpose clear	23.5	35.3	41.2	30.8	46.2	23.1	41.7	25.0	33.3	11.1	66.7	22.2	8.3	8.3	83.4
11 context time reading	11.8	52.9	35.3	23.1	53.8	23.1	33.3	25.0	41.7	11.1	11.1	77.8	41.7	33.3	25.0
12. reading time reading	0.0	41.2	58.9	30.8	53.8	15.4	0.0	58.3	41.7	33.3	0.0	66.7	0.0	41.7	58.3
13 vocabulary time	11.8	52.9	35.3	30.8	46.2	23.1	8.3	50.0	41.7	0.0	11.1	88.9	0.0	33.3	66.7
14. context time tasks	23.5	41.2	35.3	23.1	69.2	7.7	8.3	50.0	41.6	11.1	11.1	77.8	16.7	16.7	66.7
15 reading time tasks	5.9	58.8	35.3	23.1	61.5	15.4	8.3	58.3	33.3	11.1	22.2	66.7	0.0	41.7	58.3
16. results tests	17.6	76.5	5.9	38.5	53.8	7.7	33.3	33.3	33.3	22.2	22.2	55.5	8.3	41.7	50.0

TS1 = Spanish Training Programme - School 1, Class 1

TS2 = Spanish Training Programme - School 1, Class 2

TE1 = English Training Programme - School 1, Class 1

TE2 = English Training Programme - School 1, Class 2

TE3 = English Training Programme - School 1, Class 3

- = negative score (1 and 2)

± = neutral score (3)

+

= positive score (4 and 5)

Table 5. Control condition. Percentages of the frequencies of the negative, neutral and positive scores. Results by school class

	Spanish CS1 (n=15)			CS2 (n=16)			English CE1 (n=17)			CE2 (n=17)		
<i>I. Programmes</i>	-	±	+	-	±	+	-	±	+	-	±	+
1 useful programme	6.7	33.3	60.0	81.3	12.5	6.3	35.3	29.4	35.3	47.1	17.6	35.3
2 pleasant programme	20.0	26.7	53.3	87.5	12.5	0.0	41.2	41.2	17.7	47.1	23.5	29.4
3 improved comprehension	26.7	60.0	13.4	68.8	31.3	0.0	35.3	41.2	23.5	58.8	35.3	5.9
4 purpose clear	6.7	26.7	66.7	62.5	6.3	31.3	29.4	17.6	52.9	29.4	17.6	53.0
5 enough diversity	20.0	26.7	53.3	68.8	18.8	12.5	35.3	35.3	29.4	23.5	29.4	47.0
6 results tasks	6.7	53.3	40.0	62.5	6.3	31.3	58.8	17.6	23.5	29.4	17.6	52.9
7 time for reading texts	13.3	13.3	73.3	37.5	31.3	31.3	17.6	41.2	41.2	35.3	29.4	35.3
8 time for tasks	13.3	53.3	33.3	25.0	25.0	50.0	11.8	58.8	29.4	17.6	29.4	53.0
9 teaching satisfying	20.0	13.3	66.6	68.8	18.8	12.5	47.1	23.5	29.4	52.9	17.6	29.4
<i>II. Tests</i>	-	±	+	-	±	+	-	±	+	-	±	+
10 purpose clear	13.3	33.3	53.4	50.0	18.8	31.3	17.6	47.1	35.3	29.4	23.5	47.2
11 context time reading	13.3	40.0	46.7	31.3	31.3	37.5	11.8	47.1	41.2	41.2	29.4	29.4
12 reading time reading	13.3	33.3	53.3	18.8	43.8	37.5	17.6	47.1	35.3	47.1	17.6	35.3
13 vocabulary. time	0.0	26.7	73.3	25.0	31.3	43.8	23.5	23.5	52.9	35.3	23.5	41.2
14. context time tasks	13.3	46.7	40.0	25.0	25.0	50.0	11.8	29.4	58.8	47.1	23.5	29.4
15 reading. time tasks	6.7	40.0	53.3	25.0	31.3	43.3	5.9	64.7	29.4	41.2	23.5	35.3
16 results tests	6.7	60.0	33.4	31.3	37.5	31.3	52.9	23.5	23.5	29.4	23.5	47.1

CS1 = Spanish Control Programme - School 2, Class 1

CS2 = Spanish Control Programme - School 2, Class 2

CE1 = English Control Programme - School 2, Class 1

CE2 = English Control Programme - School 2, Class 2

- = negative score (1 and 2)

± = neutral score (3)

+

= positive score (4 and 5)

Appendix XVIII. Main experiment: Statistics**A. SPANISH****Spanish - Context tests****Table 1. Spanish: Context pre-test**

number of students	69	
number of items	20	
	mean	s.d.
raw score	12.14	5.39
item difficulty	.61	
reliability (Cronbach's alpha)	.91	

Table 2. Spanish: Context post-test

number of students	63	
number of items	20	
	mean	s.d.
raw score	11.75	4.26
item difficulty	.59	
reliability (Cronbach's alpha)	.84	

Spanish - Reading comprehension tests**Table 3. Spanish: Reading comprehension pre-test**

number of students	72			
number of items	34			
chance score	10.92			
	mean	s.d.	lowest	highest
raw score	20.49	4.62	7.00	31.00
item difficulty	.60	.22	.18	.93
item-rest correlation	.23	.14	-.03	.60
reliability (Cronbach's alpha)	.72			

Table 4. Spanish: Reading comprehension post-test

number of students	62			
number of items	35			
chance score	11.50			
	mean	s.d.	lowest	highest
raw score	18.58	6.04	7.00	29.00
item difficulty	.53	.16	.19	.84
item-rest correlation	.30	.22	-.22	.66
reliability (Cronbach's alpha)	.81			

Spanish - Vocabulary test**Table 5. Spanish: Vocabulary test**

number of students	63			
number of items	50			
chance score	10.00			
	mean	s.d.	lowest	highest
raw score	32.10	6.73	14.00	43.00
item difficulty	.64	.24	.08	.98
item-rest correlation	.27	.20	-.22	.60
reliability (Cronbach's alpha)	.83			

B. ENGLISH**English - Context tests****Table 6. English: Context pre-test**

number of students	65	
number of items	20	
	mean	s.d.
raw score	11.44	3.51
item difficulty	.57	
reliability coefficient (Cronbach's alpha)	.78	

Table 7. English: Context post-test

number of students	69		
number of items	20		
	mean	s.d.	
raw score	10.59	3.30	
item difficulty	.53		
reliability coefficient (Cronbach's alpha)	.72		

English - Reading comprehension tests**Table 8. English: Reading comprehension pre-test**

number of students	67			
number of items	37			
chance score	11.70			
	mean	s.d.	lowest	highest
raw score	21.73	4.80	9.00	31.00
item difficulty	.59	.19	.22	.94
item-rest correlation	.20	.15	-.14	.47
reliability (Cronbach's alpha)	.69			

Table 9. English: Reading comprehension post-test

number of students	62			
number of items	35			
chance score	11.08			
	mean	s.d.	lowest	highest
raw score	17.98	4.65	9.00	28.00
item difficulty	.51	.20	.05	.90
item-rest correlation	.19	.15	-.04	.68
reliability (Cronbach's alpha)	.68			

English - Vocabulary test**Table 10. English: Vocabulary test**

number of students	74			
number of items	50			
chance score	10.00			
	mean	s.d.	lowest	highest
raw score	29.05	6.80	15.00	45.00
item difficulty	.58	.21	.15	.96
item-rest correlation	.24	.14	-.13	.53
reliability (Cronbach's alpha)	.80			

Samenvatting

Het onderwerp van dit onderzoeksverslag is contextgebruik en begrijpend lezen in een vreemde taal. Onder het begrip 'context' wordt in deze studie verstaan de voorkennis en tekstuele informatie waarvan een lezer gebruik kan maken om de betekenis af te leiden van onbekende woorden in een tekst. Het verslag bestaat uit twee delen.

In het eerste, theoretische deel werden in hoofdstuk 1 verschillende argumenten aangevoerd om te illustreren dat lezen in een vreemde taal een belangrijke vaardigheid is. Vervolgens werd kort ingegaan op de rol van lezen in het vreemdetalenonderwijs. Tenslotte werd aandacht besteed aan het vreemdetalenonderwijs in Nederland, met name aan de gevolgen van de recente onderwijsherzieningen voor het leren en onderwijzen van leesvaardigheid.

Het tweede hoofdstuk was een samenvatting van huidige theorieën over informatieverwerking en begrijpend lezen in de moedertaal. Lezen wordt beschouwd als een interactief proces: een voortdurende wisselwerking tussen 'lagere-orde' decodeerprocessen (ook wel 'bottom-up' verwerking genoemd) en 'hogere-orde' interpretatieprocessen (ook wel 'top-down' verwerking genoemd).

In hoofdstuk 3 werd ingegaan op specifieke kenmerken van lezen in een vreemde taal. Er werd geconcludeerd dat lezen in een vreemde taal verklaard kan worden met de modellen beschreven in hoofdstuk 2. In beide gevallen is er sprake van een wisselwerking tussen 'bottom-up' en 'top-down' processen. Echter, een lezer in een vreemde taal heeft zelden hetzelfde niveau van taalvaardigheid als een ervaren moedertaallezer en zal noodzakelijkerwijs compenseren voor tekorten in taalkennis. Dit betekent dat hij of zij vaak gebruik zal maken van minder efficiënte processen, zoals het letter voor letter decoderen van onbekende woorden en het voorspellen van de inhoud van de tekst op basis van zijn of haar voorkennis over het tekstonderwerp. Aan het eind van het derde hoofdstuk werd ingegaan op het trainen van contextgebruik in het vreemdetalenonderwijs. Er werd een onderscheid gemaakt tussen voorkennis over het onderwerp van een tekst (algemene context) en de syntactische, semantische, lexicale en stilistische informatie in de tekst zelf (lokale context).

In het tweede deel van het onderzoeksverslag werden de voorbereiding, uitvoering en resultaten beschreven van een veldonderzoek naar de effecten van een trainingsprogramma in contextgebruik. In hoofdstuk 4 werd het onderzoeksplan uiteengezet. Er werden drie onderzoeksvragen geformuleerd:

1. Leidt een trainingsprogramma in contextgebruik voor beginnende leerders van Spaans als vreemde taal tot positieve effecten in zowel contextgebruik als begrijpend lezen?

2. Leidt een trainingsprogramma in contextgebruik voor half-gevoorderde leeders van Engels als vreemde taal tot andere resultaten in contextgebruik en begrijpend lezen dan een vergelijkbaar programma voor beginnende leeders Spaans?
3. Bestaat er een verband tussen vier lezer-afhankelijke variabelen – kennis van het vocabulaire van de vreemde taal, attitude tegenover de vreemde taal, zelfvertrouwen, eigen inschatting van de student van de mate waarin hij of zij gebruik maakt van de context bij het lezen in de vreemde taal - en de scores in de context- en leesvaardigheidstoetsen?

In hoofdstuk 5 werden de resultaten van het onderzoek gepresenteerd. De resultaten waren niet eenduidig. Dit betekende dat het antwoord op de eerste onderzoeksvraag negatief was: op basis van de resultaten van ons onderzoek kon niet worden aangetoond dat het trainingsprogramma voor Spaans tot betere resultaten had geleid dan het controleprogramma. De tweede onderzoeksvraag kon niet worden beantwoord. Het was, op basis van onze resultaten, niet mogelijk vast te stellen of het trainingsprogramma voor Spaans tot betere resultaten had geleid dan het trainingsprogramma voor Engels of juist omgekeerd. Twee extra analyses toonden aan dat er verschillen waren in de effecten van het trainingsprogramma op klassenniveau en zelfs op individueel niveau.

In het vijfde hoofdstuk werden ook de resultaten met betrekking tot de derde onderzoeksvraag gepresenteerd. Uit de analyses bleek dat er significante correlaties waren tussen de scores in de vocabulaire- en de contexttoetsen enerzijds, en de vocabulaire- en leesvaardigheidstoetsen, anderzijds. Tegenstrijdige resultaten werden gevonden voor het belang van de variabele 'attitude tegenover de vreemde taal': voor Spaans werd een significant verband gevonden tussen deze variabele en de scores in de context- en leestoets. Voor Engels, daarentegen, bleek 'attitude' geen rol van betekenis te hebben gespeeld in ons onderzoek. Tenslotte werd een significante samenhang aangetoond tussen de variabele 'zelfvertrouwen' en de scores in de Spaanse contexttoets.

In het zesde hoofdstuk werden de resultaten samengevat van de evaluatie van het onderzoek. Alle *studenten* die hadden meegewerkt aan het onderzoek, werd verzocht na afloop een vragenlijst in te vullen. De resultaten van deze vragenlijst toonden aan dat de studenten Spaans niet erg positief waren over het trainingsprogramma in contextgebruik. De meerderheid van de trainingsgroep voor Engels, echter, had de lessen van het trainingsprogramma leuk gevonden en vond dat het programma een nuttig instrument was om te leren de betekenis van onbekende woorden af te leiden. Zowel bij Spaans als Engels was er sprake van verschillen tussen klassen met betrekking tot de evaluatie van het onderzoek.

Bij elk van de *docenten* die bij het onderzoek betrokken waren geweest werd een uitgebreid interview afgenomen. Uit deze vraaggesprekken kwamen enkele mogelijke verklaringen voor de resultaten van het onderzoek naar voren. Een belangrijk punt was dat de studenten Spaans van de trainingsconditie steeds minder gemotiveerd waren aan het programma mee te werken, omdat ze niet het gevoel

hadden er iets van op te steken. De docenten die het Engelse trainingsprogramma hadden uitgevoerd waren positiever over het onderzoek dan de docenten die betrokken waren geweest bij het Spaanse trainingsprogramma.

In hoofdstuk 7, de discussie, werden de resultaten samengevat en besproken. Ook werd ingegaan op de implicaties van de resultaten voor het onderwijs van begrijpend lezen in een vreemde taal. Een van de belangrijkste uitkomsten van ons onderzoek is dat de waarde van het trainingsprogramma in contextgebruik verschillend was voor schoolklassen en soms zelfs voor individuele studenten. We concludeerden daarom dat het belangrijk is rekening te houden met verschillen tussen leerders, dat het zinvol is om strategie-training af te wisselen met andere onderdelen die van belang zijn voor het leren lezen in een vreemde taal (o.a. vergroten van de woordenschat) en dat het nodig is dat docenten beter voorbereid worden op het begeleiden van studenten die een leestraining volgen. Hoofdstuk 7 werd afgesloten met enkele aanbevelingen voor verder onderzoek.

Curriculum Vitae

Corine van den Brandt was born in 's-Hertogenbosch on August 11, 1962. In 1980 she received her Gymnasium A diploma from grammar school "Rijksscholen-gemeenschap Dr. Moller" in Waalwijk and went to the University of Nijmegen to study Spanish Language and Literature. In 1988 she graduated; in her thesis she reported on the role of writing in the teaching of Spanish as a foreign language. She started teaching Spanish at different institutes. In September 1989, she was appointed by the University of Nijmegen to carry out the research project this Ph.D. thesis reports on. In 1993 she was appointed Junior lecturer at the Department of Business Communication Studies and Applied Linguistics at the same university. Since 1996 she has been working as a lecturer at the same Department.

Corine van den Brandt

Context Use and Learning to Read in a Foreign Language

This study aims to assess the effects of context-use training on foreign language reading comprehension. The purpose of context-use training is to enable learners to increase their understanding of a text by making inferences on the basis of both their background knowledge and the specific textual information itself. The central question is whether or not a training programme in the use of context leads to better results than a control programme that does not include the teaching of context-use strategies. The study reports on the results of two groups of students: a group of learners of Spanish at beginner level and a group of learners of English at intermediate level.

The data revealed no conclusive evidence that the training programme in the use of context led to better results than the control programme. Results did indicate however that the effect of the training programme was different for individual school classes and even for individual learners. It turned out that the success of the training programme depended on a variety of factors. These factors include students' linguistic proficiency and motivation, and situational variables such as teaching materials, timetabling and teacher characteristics.

This book is of interest to foreign language researchers and foreign language teachers.

thesus

ISBN 90-5569-110-0

ISSN LOT series: 1389-9473

Holland Academic Graphics - The Hague

www.thesus.com