Riikka Alanen

Grammaticality judgments and reaction time measurement: a tool for analyzing the use of second language knowledge

Esitetään Jyväskylän yliopiston humanistisen tiedekunnan suostumuksella
julkisesti tarkastettavaksi yliopiston vanhassa juhlasalissa (S212)
joulukuun 20. päivänä 1997 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Humanities of the University of Jyväskylä,
in Auditorium S212 on December 20, 1997 at 12 o'clock noon.

ABSTRACT

Alanen, Riikka
Grammaticality judgments and reaction time measurement: a tool for analyzing the use of second language knowledge
Jyväskylä: University of Jyväskylä, 1997,
Yhteenveto: Kieliopillisuusarviot ja reaktioaikamittaus vieraan kielen tutkimusvälineenä
Dissertation

In recent years, there has been a lot of discussion about the suitability of grammaticality judgments for second language research. This study investigates what L2 grammaticality judgments combined with reaction time measurement can reveal of the nature of language learners' linguistic knowledge. Results form three experiments will be discussed.

In Experiment 1, RT judgments were used to obtain information about the degree to which Finnish L1 language learners relied on their English L2 knowledge in making judgmens on German L3 sentences. The experiment included two groups of German L1 learners of English, which gave an opportunity to examine the effect that different L1s had on the judgments of English L2 sentences. Fifteen NSs of English participated in the experiment as well. The NSs turned out to be significantly faster and more accurate than language learners in their judgments, which was taken to be an indication of their higher language proficiency. Evidence for language transfer was obtained as well. Moreover, consistent differences emerged in the speed of individual learners' judgments across the languages.

In Experiment 2, an attempt was made to correlate learner performance in RT judgment tasks on English L2, Swedish L3 and Finnish L1 to cognitive and affective variables including field-dependence/independence and L2 communicative anxiety. The latter turned out to have some relation to the speed of judgments. Again, the NSs of English were faster and more accurate on English than the Finns. Similar differences in individual test performance emerged as in Experiment 1.

Sixty-six first and second year Finnish L1 students of English at the University of Jyväskylä participated in Experiment 3. The aim of the experiment was to study the relationship of the speed and accuracy of English L2 grammatical judgments to L2 proficiency. Data included, among other things, the students' judgments and response times to visually presented isolated L2 English and L1 Finnish sentences, their judgments in writing on the same set sentences, and measures of L2 communicative anxiety and L2 proficiency. In addition, information was collected on various background variables. The results indicated that there was a relationship between the learners' performance in RT judgment tasks and their scores on the L2 proficiency test, especially in its grammar section. Furthermore, the cluster analysis revealed an interesting pattern in the students' performance, suggesting that there may be differences among subjects in the nature and use of L2 knowledge not immediately apparent. Methodological and theoretical implications for further study will be discussed.

Keywords: second language acquisition, grammaticality judgments, reaction time measurement, language transfer

ACKNOWLEDGMENTS

This study incorporates many years of research which began in the Jyväskylä Cross-Language Project in the mid-1980s. I am deeply grateful to all those who have given me help and encouragement during these years. Above all, I wish to thank my supervisor professor Kari Sajavaara for his unfailing support, material and nonmaterial, which has made this work possible. His help has been invaluable in bringing this work into focus. I also wish to thank professor Jaakko Lehtonen for his insight and support during the early stages of this research.

I am greatly indebted to Dr. Päivi Pietilä of the University of Turku and my external reviewer Dr. Hannele Dufva from the Center for Applied Language Studies at Jyväskylä for reviewing the manuscript of the present dissertation. Their insightful suggestions were of enormous help. I also wish to thank my other external reviewer Dr. Richard Schmidt of the University of Hawai'i at Manoa for his comments on the research and his help in shaping my thinking on research in general during my stay at Hawai'i. My sincere thanks go to professors Raija Markkanen and Liisa Lautamatti from the Department of English at Jyväskylä for their many helpful suggestions.

I also wish to thank all those people who participated in the experiments, first among them the students at the Department of English at Jyväskylä who let themselves be persuaded to give their time and effort for this study. In addition, my thanks go to all the native speakers of English and German who agreed to take part, many of them teachers at the University of Jyväskylä.

I am deeply grateful to my friends and colleagues Anne Pitkänen-Huhta, Paula Kalaja, Tarja Nikula, Sirpa Leppänen, Arja Piirainen-Marsh, Mervi Eloranta and Kaarina Mäkinen for their friendship and insightful discussions on the nature of research in general. I also wish to thank Tuula Hirvonen, Liisa Lautamatti and Raija Markkanen for their advice and many practical suggestions concerning the nature of the proceedings. Last but not least, I wish to thank my parents and brother for providing support and love and keeping my feet firmly on the ground.

LIST OF TABLES

LIST OF FIGURES

CONTENTS

1 Introduction

1.1 Orientation

There has been a great deal of discussion concerning the role of grammaticality judgments and reaction time measurement in second language research. Grammaticality judgments have been used to study language learners' linguistic competence, within the UG framework, in particular. In recent years, however, the increasing awareness of the shortcomings of this method in second language research has made researchers more careful in its use, as pointed out by Birdsong (1989) and Ellis (1991). At the same time, however, reaction time measurement has been widely utilized as a tool for investigating human performance in various types of task. A combination of the two approaches can be useful in obtaining valuable information on learner performance such as the degree of automat-ization of certain types of linguistic knowledge.

The terms grammaticality judgments and acceptability judgments need clarification. Grammaticality judgments have usually referred to grammatical competence, a speaker's knowledge of grammar as a rule system. Initially, grammaticality judgments were understood to be direct reflections of the intuitive knowledge of language, ie. competence. Acceptability is a broader term; it has generally been used to refer to the aspects of language performance, its use in social context. Acceptability judgments are therefore part of language performance. In the Jyväskylä Cross-Language Project, the two techniques of acceptability judgments and reaction time measurement were combined to study the mechanisms involved in second language processing by Sajavaara and Lehtonen in the early 1980s. Various reseachers in the United States and United Kingdom (see eg. Bley-Vroman and Masterson 1989, Cowan and Hatasa 1994, Gass 1994,) have also used reaction time measurement as a supplement to grammaticality judgments. However, despite the outward similarities of the techniques used, the aims and theoretical basis of the research approaches are somewhat different. Grammaticality judgments of L2 learners have been widely used to study L2 competence within the Chomskyan paradigm. Sajavaara and Lehtonen (1980, see also Lehtonen and Sajavaara 1985) started out from the assumption that the rate and accuracy of the reactions reflected various processual features which were connected, for example, to the occurrence of L1-like constructs in L2 speech and writing. The reaction-time method was applied to the testing of grammatical and pragmatic acceptability and to various test settings associated with lexical access, lexical decisions, and grammatical structures. The present experiments for the most part involve judgments on the grammatical acceptability of sentences, that is, the error or deviancy lies in the violation of the grammatical system. Therefore, the terms grammaticality judgment and acceptability judgment are used interchangebly (but only to refer to the present experiments), with the understanding that within the framework of the present experiments, the term grammaticality judgment always refers to judgments of grammatical acceptability. In addition, the term RT judgment will be used to refer to the grammaticality judgments for which the subjects' response time has been measured.

1.2 Aims and methods

The focus of this research is methodological. It attempts to find answers to the questions, What is the nature of this type of language task? What kind of factors are involved? and What does this type of language task tell us about the nature of language proficiency and the use of linguistic knowledge?

The data under investigation has been gathered over the past ten years in three different experiments that each involved the combination of acceptability judgments with reaction time measurement. Some of the results of the earlier experiments have been reanalyzed and thus been found to shed new light on the methodological considerations that might concern the use of this type of methodology.

In each of these experiments, English L2 learners were asked to judge sentences under circumstances in which they were susceptible to cross-linguistic influence, ie. language transfer. In other words, the sentences the students were asked to judge were designed in such a way that they might be tempted to rely on their knowledge of L1 or any other language of which they had reached a level of some automaticity. L2 knowledge, its use and the notion of automaticity are important for understanding the processes underlying the RT grammaticality judgments. Another concept that will be discussed is the interaction between the learners' systems of language knowledge, which is one of the ways cross-linguistic influence, or language transfer, can be viewed. Consequently, each of these issues merits a section of its own.

The first of the experiments to be reported here was conducted to find evidence of, first, L3 interference, that is, influence of a stronger L2 on a weaker L3, and second, influence of different L1s on English L2. The data was used for the present writer's Pro Gradu Thesis (Alanen 1987) and, later, Licentiate Thesis (1991), but has now been reevaluated and reanalyzed. When the results of this experiment were analyzed, certain interesting trends came out among the learners that seemed to point to a greater role of individual differences during the performance of this type of task. This led to Experiment 2, in which an attempt was made to correlate the learners' performance in grammaticality judgments on various languages with such factors as cognitive style (in its dimensions of field dependence/independence), thinking style (non-verbal vs verbal), L2 communicative anxiety, and ability of the verbal reasoning in L1. Partly on the basis of the results of Experiment 2, a third series of experiments was conducted that focused on the potential correlation between English learners' performance in this task type, the learners' overall English L2 proficiency, L2 communicative anxiety and their performance in a similar task in their L1, Finnish.

1.3 The outline of the study

The presents study is structured as follows: In Chapter 2, some of the issues concerning second language acquisition will be discussed, including implicit and explicit knowledge and the role of consciousness. Various cognitive theories of second language learning will be brought up such as Anderson's ACT* theory of skill learning and McLaughlin's cognitive theory of second language learning. The notions of declarative an procedural knowledge will

be discussed in more detail. Chapter 2 ends with a short description of the main points of Universal Grammar.

Some of the issues related to the use of second language knowledge will be dealt with in Chapter 3. The discussion will start from the level of language processing and go on to more general models of second language performance. The chapter ends with the description of factors involved in second language metalinguistic performance. In Chapter 4, the interaction between language knowledge systems, as language transfer can be viewed, will be discussed in some detail. Since the design of Experiments 2 and 3 involves a number of cognitive and affective variables, they will be introduced in Chapter 5. In Chapters 6 and 7, grammaticality judgments and reaction time measurement will be discussed.

The experiments included in the present thesis will be dealt with in Chapters 8, 9 and 10. Conclusions, limitations and suggestions for future research will be given in Chapter 11.

2 Second language acquisition and language knowledge

In this study, a cognitive perspective of second language acquisition will be adopted. As Ellis (1994: 347) points out, within this framework language acquistion is seen as a mental process involving the use of strategies that explain how the L2 knowledge system is developed and used in communication. Linguistic knowledge is generally regarded as no different from other types of knowledge, and the strategies involved as general in nature. Ellis (1994) contrasts this perspective with a linguistic theory of L2 acquisition (such as Universal Grammar) which sees linguistic knowledge as unique and separate from other knowledge systems, and acquisition as a result of mechanisms that are at least for the most part linguistic in nature. (The field of research can be divided differently: McLaughlin (1987), for example, divides the field in the same way into approaches that are either linguistically motivated or cognitive-psychological theories. In addition to Universal Grammar, however, McLaughlin includes Krashen's Monitor Model and Interlanguage theory among the linguistically motivated theories. Ellis (1994: 355) justifies his classification by pointing out that although Krashen relied in his theory on linguistic concepts, such as Chomsky's notion of innate linguistic knowledge, the distinction between implicit and explicit knowledge, which is central to the Monitor Theory, is a cognitive one.

Since the methodology of the present study involves grammaticality judgments, a tool used in UG-based research in particular (although by no means exclusively), it is important to point out some of the differences between the cognitive and linguistic views of second language acquisition. Although all second language acquisition research can be said in a sense to have a description of language learners' L2 knowledge as its goal, linguistic accounts are directed at describing learners' competence, seen as an abstract system of rules and items underlying actual performance. A sharp distinction is drawn between competence and performance, what learners 'know' and what they 'do', and frequently only the former is seen as worthy of study (see eg. Gregg 1990). The ultimate nature of this dichotomy has given rise to a spirited debate on the nature and development of second language acquisition theories (see eg. Eckman 1994, Brown et al. 1996).

The cognitive approaches to second language acquisition are also concerned with what learners 'know', but in their view, knowledge cannot be separated from actual use. According to Ellis (1994: 348), they focus on investigating the extent to which the learner has achieved mastery over the formal and functional properties of language and the mental processes involved. Most importantly, mastery is not seen as 'either/or' but rather considered gradable. These theories also assume that there are degrees of 'knowing'.

These two perspectives are, however, not mutually exclusive, and as Ellis (1994: 347) points out, a comprehensive theory of second language acquisition will very likely be needed to combine both these views.

Ellis (1994:) futher divides the cognitive approaches into those that attempt to explain how learners construct their mental representations of the L2, and those that explain how learners employ their knowledge in actual language use. In the following, an attempt will be made to outline some of the major approaches within this framework.

2.1 Cognitive approaches to second language acquisition

Following Ellis (1994), a general theoretical framework for cognitive accounts of second language acquisition is first presented. The concepts of implicit and explicit knowledge, and the role of consciousness will also be discussed.

2.1.1 Theoretical framework

One of the simplest ways of describing the second language acquisition process is to consider it a process whereby L2 input becomes L2 intake. In second language research, the term input is generally used to refer to the target language material the learner is exposed to during the process of language learning. The intake refers to the amount of input the learner has actually incorporated into his or her developing knowledge system, i.e. has learned. Chaudron (1985: 2) points out that intake actually involves not just one stage, or single event or product, but several stages in a complex series of information processing. Chaudron characterizes these stages as (1) the initial stage of perception of input; (2) the subsequent stages of recoding and encoding of the semantic (communicated) information into long-term memory; and (3) the series of stages by which learners fully integrate and incorporate the linguistic information in input into their developing grammars. Chaudron sees this as a continuum from preliminary intake to final intake, i.e. learning. This process as a whole is affected by the learner's state of interlanguage knowledge and the nature of constraints that operate on language perception and comprehension processes.

Gass (1988) distinguishes between four major stages: (1) apperceived (or noticed) input; (2) comprehended input; (3) intake; and (4) integration. At the first stage, features in the input are noticed as a result of their saliency and the learners' existing L2 knowledge. Not all apperceived input is comprehended, and not all comprehended input becomes intake. Gass (1988: 206), following Chaudron (1985), sees intake as a process mediating between L2 input and the learner's internalized set of rules, ie. his or her interlanguage system. Intake can also been seen as the end product of the process of intake: VanPatten (1989, see also 1990), for example, defines intake as a subset of the input that the learner actually perceives and processes. It does not, however, become part of the learner's implicit knowledge system until it has been incorporated or integrated.

To these stages, Ellis (1994: 349) adds the component of explicit knowledge (shown below in Figure 1). Thus, he uses Gass's (1988: 207)) suggestion that some input may be processed and 'put into storage' if it cannot yet be integrated into the interlanguage system. This storage, according to Ellis, can take the form of some kind of explicit representation of L2 items and rules. Explicit knowledge can contribute to output through monitoring, and also may aid the processes that contribute to intake. Finally, output can influence intake through interaction.

The theoretical positions discussed in this chapter address different aspects of this framework. Thus, the notion of interlanguage and interlanguage theory (Selinker 1972) is primarily concerned with implicit L2 knowledge and its development. Krashen's Monitor Theory and Bialystok's bidimensional model of second language acquisition are mainly concerned with identifying the relationship between implicit and explicit knowledge and how they are used in L2 output. Anderson's ACT* theory and McLaughlin's information processing model are based on skill-learning. According to these views, the use of L2 in

tasks that place different demands on the learner's processing abilities is linked to the changes in the representation of knowledge.

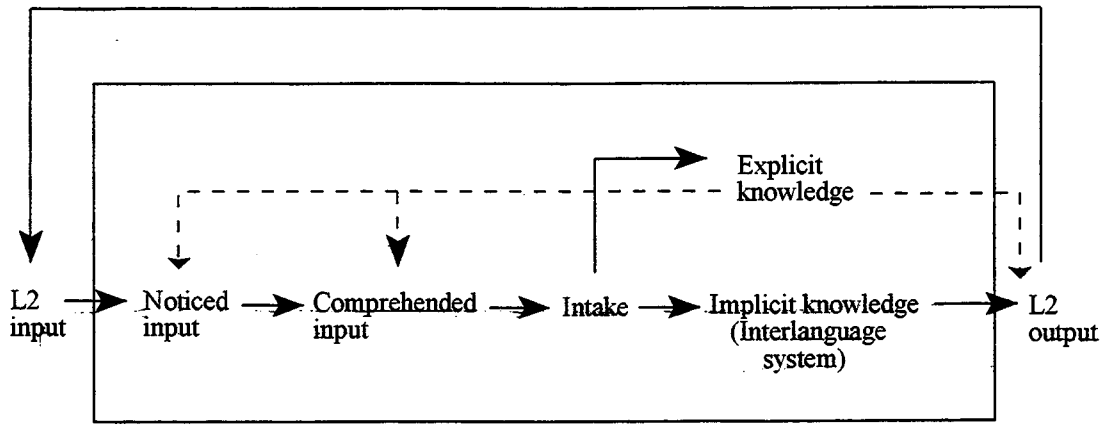


FIGURE 1 A framework for investigating L2 acquisition (from Ellis 1994: 349).

2.1.2 Interlanguage theory

The emergence of the notion of interlanguage in the early 1970s changed the way of thinking about language learning. Similar ideas had been presented a little earlier by Corder (1967), who talked about 'transitional competence', and Nemser (1971), who discussed the language learner's 'approximative systems'. Interlanguage theory (Selinker 1972), however, was the first major attempt to provide an explanation for L2 acquisition. Selinker first introduced the term interlanguage in 1969 (Selinker 1992: 231), but the seminal article in which he identified five cognitive processes he saw responsible for second language acquisition appeared in 1972.

Interlanguage is quite often seen as a distinct linguistic system which is the result of the learner's attempts to achieve target language norms. The underlying assumption is that the system has a relative stability, and that there are patterns and features typical of its various stages. Interlanguage contains deviations from the target language norm, that is, errors, the number and nature of which can be seen as changing as the learner makes progress to higher levels of language proficiency. The learner has then a competence in his second language, and as for example Schachter, Tyson and Diffley (1976: 69) point out, the learner can also then be asked about his or her intuitions about interlanguage. However, it is important to note that the term interlanguage has other meanings, as well. First, it can be used to refer to the learner's system at a single point of time or a series of connected systems that the learner develops over time (McLaughlin 1987: 60). Secondly, it can be used to refer either to the product or to the underlying processes that have led to that product (Sharwood Smith 1994: 35). Thirdly, interlanguage can be defined as "the systematic linguistic bahaviour of second language learners" (Sharwood Smith 1994: 6). This is in direct contrast to many other researchers as Sharwood Smith (1994: 7-8) makes a point in using interlanguage to refer to "the actual observable linguistic behaviour of language learners" (1994: 35), rather than the 'invisible' language system.

According to Selinker (1972: 116f.), who was the first to use this term, interlanguage (or to use James's term (1980: 160), "interlingua"), derives its features from five processes:

1. Language transfer. Some, but not all, items, rules and subsystems of a learner's interlanguage may be transferred from the first language.

2. Transfer of training. Some interlanguage elements may derive from the way in which the learners were taught. Selinker (1972: 218) gives an example how the mode of learning in a classroom context can influence what learners transfer: the fact that a Serbo-Croatian learner of English uses the personal pronoun *he* in cases where he should use *she* even though his L1 makes a corresponding distinction between feminine and masculine, can be explained by the material and the method (especially if it is of the type that uses drills) applied in language teaching.

3. Strategies of second language learning.

4. Strategies of second language communication.

5. Overgeneralization of the target language material. Some interlanguage elements are the result of overgeneralization of L2 rules and semantic features.

Selinker, Swain and Dumas (1975) later specified this early view: they now saw interlanguage as a system the rules of which had been developed through the application of various cognitive strategies such as transfer, simplification and overgeneralization. The first and the latter two of the processes mentioned above have often been seen as 'competing' phenomena: both have been assigned the central role as the relevant strategy of learning (cf. above; the conflict between the advocates of morpheme studies and language transfer). They have also been combined under the same concept: thus Arabski (1968: 73-74), for example, calls overgeneralisation "internal interference, called also analogy", and transfer "external interference"[1]. It may well be, as R. Andersen (1983: 177) points out, that treating transfer and overgeneralisation as two distinct, mutually exclusive forces may cause us to disregard the way transfer determines the learner's perception of structural relationships within input.

As Gass and Selinker (1983: 7) point out, today it has become possible to see second language learning both as a process of hypothesis testing in which learners form their second language knowledge on the basis of second language data available to them, and as a process in which learners utilize their L1 knowledge as well as knowledge of other languages in order to create a learner language, or interlanguage. When the learner forms hypotheses about the target language (TL) on the basis of input which is interpreted by means of the strategies utilising the learner's L1 (or other languages different from the relevant non-native language), such formation process is known as transfer (Faerch, Haastrup and Phillipson 1984: 192).[2]

Second language learning can thus be seen as a primarily cognitive process of hypothesis formation and hypothesis testing. Closely linked to this view is the notion of

---

[1] There are certain similarities between these two strategies; both seem to utilize a certain amount of metalinguistic knowledge whether it is about L1 and its relation to L2, or L2, or combination of these both.

[2] According to Faerch, Haastrup and Phillipson (1984: 192), there are two other possible strategies that the learner might use; in addition to overgeneralisation (they use the term generalisation), in which the learner resorts to his knowledge of his existing interlanguage, she may also apply productive strategies known as *communicative strategies* which also involve the activation of L1 and second language.

control. Learners do not have to be aware of these hypotheses; on the contrary, learners often seem to be unconscious of them. Thus, such hypotheses are rather to be seen as cognitive representations of the learner's interlanguage knowledge, not conscious beliefs.

The great significance of interlanguage theory was that it started to shift the focus away form product to process in second language acquisition studies. Much of the subsequent work has focused on either learning strategies or, at a more general level, the notion of hypothesis-testing. Hypothesis-testing refers to the process whereby learners make hypotheses about the target language structure on the basis of the input data they are exposed to, which they then test both receptively and productively. All these concepts have been more or less incorporated in the terminology of second language acquisition studies. The definitions of these terms as well as of those included by Selinker in his interlanguage theory (such as the relationship of a strategy of learning to overgeneralization) remain, however, rather vague.

2. 1.3 Implicit/explicit knowledge and the role of consciousness in second language acquisition

The definition of implicit and explicit knowledge

Linguistic knowledge is traditionally described as rules and items. (It can also be regarded as a network involving a complex set of interconnections between various units as in Parallel Distributed Processing theory (see Ellis 1994: 348).) Explicit knowledge generally refers to the knowledge available to the learner as a conscious representation. It is not the same as formal metalinguistic knowledge, which can be defined as the knowledge of the special terminology used for linguistic concepts. On the contrary, learners may make their knowledge explicit in everyday language. There are two types of implicit knowledge, formulaic knowledge and rule-based knowledge. Formulaic implicit knowledge consists of language chunks. Rule-based implicit knowledge consists of generalized and abstract structures which have been internalized. Implicit knowledge is intuitive in nature and learners are not conscious of what they know. It becomes manifest only in actual performance.

Implicit and explicit knowledge can also be described in the following way: implicit knowledge is what the individual language speaker can use but not describe, whereas explicit linguistic knowledge is knowledge which the individual can describe but does not necessarily use (Faerch, Haastrup and Phillipson 1984: 201). The notion of rule is not to be understood as equivalent of explicit rules of grammar but seen as related to notions such as "processing strategies", "knowledge of underlying structure" (Slobin 1979: 5; 98). Rather than to consider these two concepts a dichotomy, Faerch, Haastrup and Phillipson (1984: 202) propose a continuum (shown in Figure 2) where the different types of linguistic knowledge are characterised in terms of consciousness.

The reason given by Faerch, Haastrup and Phillipson for positing the two left-most categories as separate is that there is no one-to-one correspondence between actual language use and acceptability judgments; the awareness of the existing linguistic norms may affect the individuals' decisions as well.

The role of explicit and implicit linguistic knowledge has been seen to be of great importance in second language learning and teaching (Bialystok 1978, 1979; Krashen 1981, 1982; Faerch, Haastrup and Phillips 1984; Bialystok & Sharwod Smith 1985; Sharwood

Smith 1993, 1994). This distinction is especially relevant to formal language learning situations such as second language classrooms. It is to be noted that implicit knowledge does not belong to the domain of performance but is seen as part of language competence. This includes learners' implicit knowledge of grammar, the rules of use and convention, and the rules of discourse. This notion of competence contains both grammatical and pragmatic knowledge of language and has become to be known as communicative competence.

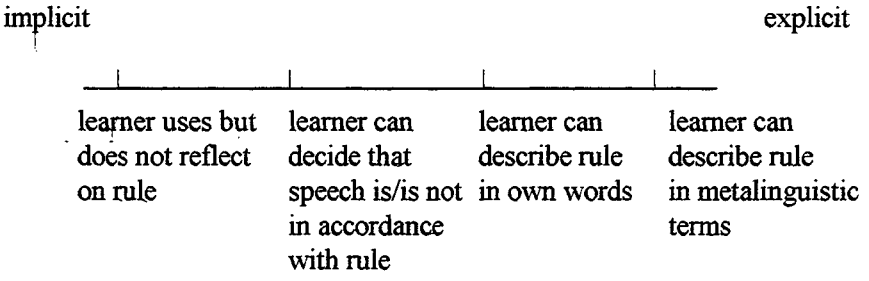


FIGURE 2 The implicit/explicit knowledge continuum (based on Fearch, Haastrup and Phillipson 1984: 202).

There are individual differences among language learners as regards their use of explicit grammatical knowledge. For example, to Krashen (1981: 15-18), such knowledge is only available for the language learner as a monitor; thus, he is able to distinguish between monitor 'overusers' who resort to formally learned rules all the time and Monitor 'underusers' who do not use conscious grammar at all. There is a third group he calls 'optimal' users. Monitoring as a psychological phenomenon is, of course, not tied to Krashen's theories

There is an interaction between the type of knowledge and consciousness: implicit and explicit knowledge form a continuum qualified by the degree of consciousness. Implicit knowledge implies low level of consciousness whereas explicit knowledge and the ability to verbalize it, ie. high level of consciousness, represent the other extreme. Two other types of knowledge can also be defined by their relationship to implicit/explicit distinction: procedural and declarative knowledge. Procedural knowledge is located at the implicit end of the continuum. It is to be noted, however, that declarative knowledge can be implicit in the sense that language learners do not necessarily have to be aware of it (eg. semantic information contained within a lexical entry in the mental lexicon); it can be made explicit, though. Indeed, it is important to keep these two dimensions of knowledge, declarative/procedural and implicit/explicit, apart from each other.

There are basically two ways of dealing with the nature and application of these two types of linguistic knowledge. For example, Krashen (1981, 1982) represents the view that implicit knowledge is completely different in nature from metalinguistic knowledge. (Interestingly, a certain amount of metalinguistic/metacommunicative awareness has been posited as one of the factors which affect the occurrence of language transfer.)

According to a different view, learning can proceed either from implicit knowledge to more explicit knowledge, or in the opposite direction. It follows that a certain rule is not restricted to one type of knowledge alone, as Krashen's Monitor Theory proposes, but that it can exist at both ends of the continuum thus formed. The learner can become more conscious about implicit knowledge (part of the general process of consciousness-raising (see eg. Sharwood Smith 1981, Rutherford 1988; Sharwood Smith (1993, 1994) now uses the term input enhancement). The learner is also able to gradually develop an ability to use a certain rule in production and reception without being aware of this. Humans may become capable

of establishing automatic routines on the basis of explicit knowledge for example in second language learning (Hulstijn 1990: 34). This process is usually referred to as an automatization process. As Faerch, Haastrup and Phillipson (1984: 203) note, such an approach has also been proposed by some Soviet psycholinguists (see eg. Leontev 1973: 12,15).

Krashen's Monitor Theory

A distinction is often made between acquisition and learning in second language research. The best-known proponent of this dichotomy is probably Stephen Krashen (1981, 1982). According to Krashen (1982), adults have two ways of developing ability in second languages. Second language acquisition is a subconscious process very similar to the process children use in acquiring first and second languages. Second language learning, however, is a conscious process; it is thought, in contrast to the former, to be greatly helped by such techniques as error correction and the presentation of explicit rules. As McLaughlin (1987: 20) points out, it is not so much the setting as conscious attention to rules that distinguishes language acquisition from language learning; acquisition, after all, is achieved without conscious focusing on linguistic forms (Krashen 1978).

The theoretical framework proposed by Krashen, known as the 'Monitor Model'or 'Monitor Theory', has been severely criticized as it leaves many of the key concepts undefined (see eg. Sharwood Smith 1981, Gregg 1984, McLaughlin 1987). According to Sajavaara (1987a), one of the main problems with Krashen's theory is that it confuses knowledge with processes. VanPatten (1994) makes a similar point in discussing Krashen's view of consciousness. On the one hand, Krashen makes a distinction between competence, subconscious, implicit, informal knowledge about the structures and rules of a language, and acquisition, a process whereby this knowledge is acquired. On the other hand, learning is referred to as product, instead of process, related to conscious, explicit, formal knowledge about the rules of the language. The claim that there are two types of knowledge, implicit and explicit, is not controversial (see eg. Ellis 1994: 356). Furthermore, as Johnson (1996: 50-51) points out, the existence of these two different processes is not disputed by many researchers. Many reseachers, however, find untenable Krashen's insistence that 'learning ' and 'learned' knowledge is a totally separate system from 'acquisition' and 'acquired' knowledge, a position known as the non-interface position (as opposed to the one known as interface position that allows for interaction between the two systems).

In the present study, the term second language learning is adopted without making any distinction between the two types of learning; it is to be thought as a general term covering both modes. If need arises, the distinction will be made and commented upon. It is an interesting question whether the language systems thus learned show any differences; are there any differences in performance, or are the automatized processes of an 'acquired' language any different from those of a 'learned' language? It seems likely that most language learners learn their language both explicitly and implicitly. Whether this difference between two learning modes has any bearing on the functioning of transfer is not clear either.

Bialystok's bidimensional model

Bialystok's (1978; see also 1979) theory of second language learning was originally based on the same dichotomy between implicit and explicit knowledge as Krashen's Monitor Theory was. According to Bialystok (1978), implicit knowledge is developed through exposure to communicative language use and it is facilitated by fucntional practising. Explicit knowledge

arises when learners focus on the language code and is facilitated by what Bialystok calls formal practising, which involves either conscious study of the L2 or attempts to automatize already learn explicit knowledge. Unlike Krashen's theory, Bialystok's 1978 model allows for interaction between the two types of knowledge, as explicit knowledge may become implicit through formal practising and implicit knowledge can serve as the basis for explicit knowledge by inferencing. Bialystok's 1978 model also has two different types of output, one which is spontaneous and immediate and relies on implicit knowledge, and the other which is deliberate and occurs after a delay and involves both implicit and explicit knowledge.

Bialystok (1981, 1982, 1990a, 1991; see also Bialystok and Ryan 1985a, 1985b and Bialystok and Sharwood Smith 1985) has revised her model several times. Her later views all involve two separate dimensions for the description of L2 proficiency. Bialystok's 1990a system has, first, the dimension of 'analysis', where information can be either analyzed or nonanalyzed, the former being the case when the learner could become explicitly aware of the structure (see below, however), and secondly, the dimension of 'control', where production or application of procedures can be either automatic or nonautomatic.

The notion of control has been connected in one form or another to various theories of second (foreign) language learning and processing. One such attempt was made in the theoretical framework of Krashen's Monitor Theory where an attempt was made to explain learner performance through the notion of monitor. Explicit knowledge or knowledge that had been obtained through 'learning' was only available for the L2 user through 'Monitor'. There is also a more general sense for monitoring: L2 learners vary to the degree their monitor their speech (see eg. Hulstijn 1982).

Bialystok's definition of 'control' has changed somewhat over the years. Bialystok (1982) used 'control' to refer to the ease of processing: how fast and easily knowledge could be accessed in different types of language use. In Bialystok and Ryan's (1985a, 1985b) theory of metalinguistic skill that included both dimensions of knowledge and cognitive control, 'control' is still used to refer to the ease of processing but in a more refined way; 'control' now involves three different functions: the selection of the items of knowledge, their co-ordinatation, and the extent to which selection and co-ordination can be carried out automatically. Moreover, in their view, the development of automaticity does not imply change in the nature of knowledge from declarative to procedural as in Anderson's model (Bialystok and Ryan 1985b: 235).

Bialystok (1990a, 1991) refers by 'analysis' to 'analysis of knowledge'; in fact, in her earlier models (1982, for example), she called her two dimensions 'knowledge' and 'control'. She sees 'analysis' as a process by which mental representations of linguistic knowledge are built up, structured and made explicit for the learner. This process starts with unanalyzed knowledge, which the learner then goes on to develop into analyzed knowledge. As Ellis (1994: 358) points out, it is tempting to see the dimension of 'analysis' as equivalent to the implicit/explicit distinction, with unanalyzed knowledge corresponding to implicit knowledge and analyzed knowledge corresponding to explicit knowledge. In Bialystok's view, however, analyzed knowledge need not involve consciousness. This contrasts with the traditional definition of explicit knowledge as the kind of knowledge available to learners as conscious representations.

Schmidt's view on the role of consciousness in second language acquisition

One of the fundamental questions not only in second language research but also in learning psychology in general has been: Are learners aware of what they are learning? Or, do learners

need to be aware of what they are learning in order for learning to take place? The positions adopted in second language research vary from the basic assumption that all language acquisition is fundamentally unconscious (e.g. Krashen 1981) to the claim that attention is the necessary and sufficient condition for converting input to intake (see e.g. Schmidt 1990, 1992, 1994; see also Schmidt & Frota 1986; Chaudron 1985).

The two major problems in the discussions on and research in the role of consciousness in second language acquisition appear to be the vagueness with which the term consciousness is used and its fundamentally subjective experience (Schmidt 1994: 12). For example, McLaughlin (1990) has recommended that the terms conscious and unconscious should be abandoned, as they are, in his view, 'pre-scientific', and furthermore, refer to subjective experience. McLaughlin (1990) reviews the history of conscious and unconscious processes as they have been applied in the study of second language acquisition. He distinguishes between three great debates:

1. The Krashen/McLaughlin debate. In contrast to Krashen's acquisition/learning distinction, McLaughlin argued for a distinction between controlled and automatic processes. As he points out, this distinction is not to be equated with conscious/unconscious distinction (see eg. Shiffrin and Schneider 1977). The distinction only relates to the the degree to which a cognitive skill has been routinized and established in long-term memory, which is empirically testable. (On the other hand, McLaughlin does not mention the potential problem of empirically differentiating between automatic and speeded-up, 'veiled' controlled processes (Shiffrin and Schneider 1977: 138; see eg. Segalowitz and Segalowitz 1993).)
2. The Reber/Dulany debate. Reber (1976, 1989) and his coworkers have argued for implicit learning of abstract rule-systems. Their findings have been argued against by Dulany, Carlson and Dewey (1984).
3. The McClelland, Rumelhart/Pinker, Prince debate.The question here is whether knowledge of language consists of mentally-represented rules (argued for by Pinker and Prince 1988) or whether it is encoded in a network of units or nodes receiving activity along the connecting arcs (a position argued for by McClelland and Rumelhart 1985).

The main argument McLaughlin seems to be making in each case is the question of empirical testing: for example, Krashen's view of unconsciousness/consciousness is not falsifiable while his is. While acknowledging Schmidt's (1990) view of the problematic areas in the concept of consciousness, he advises caution in using the term.

Schmidt (1994), partly in response to McLaughlin's arguments, outlines a number of concrete proposals for further study. He distinguishes between four senses of consciousness:

1. Consciousness as intentionality. According to Schmidt (1994: 16), an important issue in second language learning is whether the knowledge gained through incidental learning has a different mental representation from knowledge gained through intentional approaches to learning. Incidental learning, ie. learning without volition, in itself is clearly possible.
2. Consciousness as attention. Consciousness can be regarded as the product of an attention mechanism. 'Focal attention' is sometimes referred to as consciousness. It can also be argued that learning without attention is impossible.
3. Consciousness as awareness. The issue of implicit and explicit knowledge and learning falls within this category.
4. Consciousness as control. The fundamental question connected with control is the development of automaticity and fluent performance.

Schmidt (1990) distinguishes between three key issues that are related to the role of consciousness in input processing:

1. subliminal learning, by which Schmidt refers to the question whether conscious awareness

at the level of 'noticing' is necessary for language learning;
2. incidental learning, which refers to the question whether it is necessary to consciously 'pay attention' in order to learn; and
3. implicit learning, which Schmidt uses to refer to the question whether learner hypotheses based on input are the result of conscious insight and understanding or an unconscious process of abstraction. In research literature, there is a considerable amount of ambiguity about the use of these terms: thus, Hulstijn (1989) sees lack of intention as a major characteristic of incidental learning, in which case consciousness can be identified with intention (cf. Schmidt 1990: 133).

Schmidt (1990) proceeded from the initial hypothesis that noticing is the necessary and sufficient condition for converting input to intake; he later modified his view by hypothesizing that attention is a sufficient condition for learning (Schmidt 1994) As used by Schmidt (1990), noticing refers to what is also known as focal awareness. Above all, noticing is a private experience, although it can be operationally defined as availability to verbal report, with certain reservations: first, the lack of a verbal report cannot be taken as evidence of failure to notice unless the report is gathered either concurrently or immediately following the experience; secondly, it has to be remembered that there are conscious experiences that are inherently difficult to describe. If these reservations are kept in mind, verbal reports can be used to verify, or falsify, claims concerning the role of noticing in cognition.

Schmidt and Frota (1986) constitutes an example of how such verbal reports can be used to gain insight into the effect of noticing. The subject, who was Schmidt himself, described his experiences in learning Portuguese as a second language in Brazil. One instance in particular seemed to constitute an example of implicit learning: it appeared that the subject had actually learned how to use aspect without being aware of it. However, further analysis of the data revealed that the use of the aspect was linked to particular verbs, which suggested that the subject had used a lexical strategy for aspect. When individual verbs were analyzed, there appeared to be a connection between the inherent lexical semantics of the verbs and the choice of aspect: thus, the subject used stative verbs in the imperfect and nonstatives in the perfect. However, there was also a high correlation with input frequency, which seemed to suggest that the subject's productions were influenced by the greater noticeability of such verbs.

Schmidt (1990: 143) comments on several factors that seem to affect the noticeability of elements in linguistic input:
1. Expectations. Events that cannot be interpreted in the current context or which are so stable as to be part of the context remain unconscious. In second language acquisition, innate universals and expectancies based on the knowledge of native and target language systems may constrain noticing without the language learner ever becoming aware of it. It is also quite possible that instruction may affect noticing of features in input by creating such expectancies.
2. Frequency.
3. Perceptual salience. This is seen as a major factor affecting L1 acquisition by Slobin (1985: 1164), who reports that children find it more difficult to acquire grammatical morphemes that are bound, asyllabic, unstressed, or varying in form.
4. Skill level such as the degree of automaticity in language processing may affect noticeability, as well. For example, the acquisition of a new syntactic structure in L1 development may depend on the extent to which previous structures have been routinized.
5. One of the most powerful determinants of what is noticed are task demands (see eg. Ericsson and Simon 1984). They also serve as one of the most important arguments for seeing noticing as the sufficient and necessary condition for learning. As the task forces the material to be processed in a certain way, it also affects what is noticed.

Schmidt has pointed out that noticing is the subjective manifestation of attention, and that it is attention that is the necessary and sufficient condition for storage in memory. (Successful storage in memory may not guarantee efficient retrieval, though (see eg. Logan 1988).) Such a view of the role of attention will necessarily have far-reaching consequences for the role attributed to second/foreign language instruction.

2.1.4. Skill-learning models of second language acquisition

Anderson's ACT model*

The distinction between declarative and procedural knowledge in its application to the area of language learning has been widely discussed by Anderson (1976,1980,1983). Anderson's ACT* (pronounced as 'star') learning theory (the term ACT comes from Adaptive Control of Thought) sees the process of language learning, whether L1 or L2, to be basically similar to the processes involved in the learning of any complex cognitive skill. Thus, Anderson undoubtedly represents one extreme within the cognitive approach to language learning whereas Chomskyan approaches, notably Universal Grammar, represents the other extreme. This aspect of Anderson's approach has been criticized to a certain extent by researchers in the field of language learning as overlooking for example the influence that the acquistion mode may have on the nature of L1 and L2 productions (Möhle and Raupach 1989: 214). The nature and primacy of declarative and procedural of knowledge is also still under debate (Möhle and Raupach 1989: 195-196). This question will ultimately be of utmost importance in the design of new theories on second language acquisition that make use of this distinction (cf. connectionist approaches).

The distinction between declarative and procedural knowledge is basically that of 'knowing that' and 'knowing how' (Ryle 1949). For example, lexical entries are typically seen as a form of declarative knowledge stored in the mental lexicon. The structure of the mental lexicon is not quite clear, however. Furthermore, in bilingual or multilingual speakers it may be that L1 and L2 lexical entries are stored in separate mental lexicons (see eg. de Bot 1992). It can be assumed that each entry contains four types of information, meaning (conceptual structure) and syntactic information (which both together can be called *lemma* information), and morphological and phonological information (known as *form* information) (see eg. Levelt 1989). All this knowledge, together with the language learners' knowledge of the world, situation knowledge, discourse record, all his various schemata and frameworks, are stored in long term memory. All this type of knowledge is called declarative knowledge.

According to Anderson, second language learning, as indeed any kind of learning of highly automatized skills, takes place when declarative knowledge is slowly transformed into procedural format, eg. through practice. This notion is useful, but it ought to be borne in mind that it is not clear how exactly declarative knowledge is transformed into some other kind of knowledge as the nature of these representations is also unclear. However, certain 'working hypotheses' can be made on the nature of mental represention of linguistic and other types of knowledge. For example, lexical entries in the mental lexicon are part of the store of declarative knowledge, and therefore it is to be assumed that when we learn new words in a second language, these are also stored in long term memory. Their use and access may or may not become more automatized as the links that connect them to our conceptual store (while at the same time reshaping it) strengthen. It is these that form a new kind of procedural knowledge. According to Faerch and Kasper (eg. 1987: 115), procedural knowledge is free of

(linguistic) content. Declarative knowledge is also changing and probably reshaped by procedural knowledge to provide easier access.

Anderson (1976: 117) makes the following assumptions on the nature of both types of knowledge: first, declarative knowledge seems, contrary to procedural knowledge, to be possessed in an all-or-none manner. However, there is a distinction between what the learner possesses and what he can access within the limitations of human performance; declarative knowledge is not stable, in a static state, as the individual is constantly editing and reorganizing it ('restructuring') (cf. McLeod and McLaughlin 1986; Karmiloff-Smith 1986a). The second assumption concerns the manner of acquisition: declarative knowledge is acquired suddenly, by being told, whereas one acquires procedural knowledge gradually, by performing the skill. (This notion has remained rather controversial.) Thirdly, one can communicate verbally one's declarative knowledge but not one's procedural knowledge. Thus, declarative knowledge can be viewed as located at the explicit end of the implicit-explicit knowledge continuum, while procedural knowledge is located at the implicit end of the implicit-explicit knowledge continuum. However, one can certainly possess declarative knowledge tacitly, without verbalizing it, until such verbalization is required. Thus, declarative knowledge can be implicit in the sense that language learners do not necessarily have to be aware of it (eg.semantic information contained within a lexical entry in the mental lexicon); it can be made explicit through verbalization. Verbalization implies the highest level of consciousness (cf. P. Andersen 1986, who speaks of the level of 'articulate' consciousness as a level one step higher than the level of 'awareness'; see also Schmidt 1990.) Anderson (1976) also proposes the existence of three types of memory, working memory, declarative memory, and production memory, which contains so-called productions. Such productions can be seen as condition-action pairs (IF-THEN pairs).

The acquisition of cognitive skills is a gradual shift from a declarative stage to a procedural stage. According to Anderson (1983: 34), all knowledge is initially declarative and must be interpreted by general procedures. During the learning process, declarative knowledge, which comes in cognitive units such as propositions (eg. *hate, Bill, Fred*), strings (*one, two, three*) or spatial images, is gradually proceduralized, composed and compiled to create task-specific routines (or productions). Anderson (1980: 222-3) distinguishes between three stages in the acquisition of procedural knowledge: first, there is an intructional or study phase called 'cognitive', 'interpretive', or 'declarative' phase, which results in "an internal and probably declarative representation of what the learner must do". The following stage, called 'associative' stage, involves the determination of the methods for performing the skill (ie. procedural knowledge), during which the declarative information is transformed into a procedural form. However, Anderson adds that both types of knowledge can exist side by side.[3] Underlying this conversion is the mechanism called knowledge compilation, which, again, can be divided into two subprocesses, proceduralization and composition. These two subprocesses seem to be conceptually similar to what McLaughlin, Rossman and McLeod (1983) and Karmiloff Smith (1986a, 1986b) call automatization and restructuring. Lastly, the learner can reach the 'autonomous' stage, in which the procedure becomes more and more automated and the ability to verbalize knowledge of the skill can be lost. (Note that Möhle and Raupach (1989: 199) at least at this point seem to be using the term procedure rather loosely, to refer to procedural knowledge.)

---

[3] In many ways, this can be seen reflected in the linguistic performance of (advanced) language learners, especially if they have learned the language in question in a classroom situation.

Declarative and procedural knowledge and the way procedural knowledge is seen to be acquired are connected to the notions of automatic and controlled processes (Anderson 1983: 34). It is in the area of access that various degrees of automaticity come into play. Procedural knowledge, which according to Faerch and Kasper (1987: 115) consists of the sum of the procedures that are used in accessing declarative knowledge for the purposes of speech reception, production, and language learning, can be employed to a varying degree of automaticity and level of consciousness.

The notions of procedural and declarative knowledge as developed by Anderson (1976,1983) have been adapted to theories of second language learning by Faerch and Kasper (1986, 1987), and Dechert and Raupach and their coworkers (see Dechert, Möhle and Raupach 1984, Dechert 1987, Möhle and Raupach 1989). This latter group of reseachers at Kassel University have investigated how learners plan their speech over time. They have focused their investigation on two phenomena: temporal variables, such as speech rate and pause length, and hesitation phenomena, such as repetitions and corrections. Faerch and Kasper's (1986,1987; see also Faerch, Haastrup and Phillipson 1984) approach is of particular interest in understanding the different aspects of language transfer. Hulstijn (1989, 1990) has also applied Anderson's notions of procedural and declarative knowledge in second language learning.

McLaughlin's cognitive theory

Central to many cognitive theories of second language learning is the idea of the language learner as a human information processor. Such an approach to language learning has close ties to language processing and inevitably also to the various models that have been developed to explain and understand language faculty in human beings. The cognitive theory of second language learning as described by McLaughlin (McLaughlin 1978, 1980, 1987; McLaughlin, Rossman and McLeod 1983) has been very much inspired by cognitive psychology and has incorporated various ideas from the field such as the notion of controlled and automatic processing

In McLaughlin's view, the learner is seen as an information processor with limits set by both the nature of the task to be carried out and his or her own information processing ability. Learners are not capable of attending to all the features of the input: instead, some of it becomes the object of focused or selective attention, while other parts are attended to only peripherally (cf. Schmidt 1990). In order to be able to process information more efficiently, learners do, what McLaughlin calls, routinization of skills. The mastery of a skill gradually develops from controlled to automatic processing, with the result that the burden on their information processing capacity is reduced.

The distinction between automatic and controlled processes is seen in the way automatic processes are regarded as quick and effortless and requiring little or no attention, while controlled processes are slow and require attention (eg. Shiffrin and Schneider 1977). The notion itself as it is commonly understood has been modified and partly criticized by Logan (1985, 1988). The mastery of any complex cognitive skill, such as second language performance consists of the skillful combination of these two processing types, as during the course of learning, the learner masters the various sub-skills and can thus devote more attention to higher-level planning or metalinguistic and metacognitive analysis required by the task. There is a useful distinction to be made between skill and automaticity, which can also be applied within this study: According to Logan (1985: 378-379), to make a clear distinction between the automatic and controlled modes of information processing implies that automatic processes can interfere with performance by typically increasing reaction time when the

decisions or responses required by the task come into conflict with them, since automatic processes are seen as difficult to control. However, it seems that automatic processes can be controlled as skilled performers are usually able to control their performance better than unskilled performers, even though their performance still retains features typical of automatic processes. Thus, good language proficiency and a large degree of automaticity in a foreign language performance usually go together.

Learners' information processing capacity is also extended by a change in the representation of knowledge. McLaughlin calls this process 'restructuring'. This involves a qualitative change in a learner's language knowledge. McLaughlin draws heavily on Karmiloff-Smith's (1986a, 1986b) views on restructuring. Karmiloff-Smith (1986a) argued that learners solve problems by going through the same recurrent phases. Phase 1 is data-driven: components of the task are mastered, but there is no attempt at overall organization. At Phase 2, organization is imposed as learners attempt to simplify, unify and gain control over the internal representation. Phase 3 involves the integration of the data-driven, bottom-up processes that guide Phase 1 and the internally-generated, top-down processes that guide Phase 2. This integration is a result of the restructuring that took place at Phase 2. By including the restructuring dimension in his theory of second language learning, McLaughlin (1987) indicates that learning involves something more than the development of automaticity.

2.1.5 The Competition Model

The Competion Model is a performance model based on a functionalist view of language. The term functionalism is generally used to refer to a wide range of approaches in psychology, philosophy and linguistics, which has led to some confusion. Bates and MacWhinney (1989: 3) define linguistic functionalism as the belief that "the forms of natural languages are created, governed, constrained, acquired and used in the service of communicative functions". In other words, they view language as a goal-directed system of activities. The model was primarily developed for the crosslinguistic study of language use by Bates and MacWhinney (1979, 1982, 1989). As Bates and MacWhinney (1989) point out, one of the most important criteria for the evaluation of this model is how well it can account for processing and acquisition in any given language.

The Competition Model rests on four tenets. First, cognition is the basis for language universals. Based on the nature of human cognition, there is "a universal set of processing constraints that sharply delimit the way that meanings and intentions can be mapped onto a real-time stream of gestures and/or sounds" (Bates and MacWhinney 1989: 6). Thus, according to the Competition Model, Chomsky's Universal Grammar is simply a manifestation of such processing constraints. Secondly, grammars are regarded as solutions to the problem of how meanings are mapped to form. Grammars are formed as a result of the interaction between cognition and constraints on human information processing. Within this interaction, there is constant competition between various processing constraints such as between the tendency to the least effort and clarity. Thirdly, although Bates and MacWhinney (1989: 10) see grammar as a biological system, language universals are only indirectly innate; they are based on interactions among innate categories and processes that are not specific to language. Fourthly, the Competition Model considers quantitative analyses as ways of understanding qualitative variation. Languages differ qualitatively (for example, in some languages there are case markings, in others they are missing) and quantitatively (even if two languages have the 'same' linguistic device, they may differ as to the extent the device is used and in the range of

functional roles it has). For example, word order and the morphological device of subject-verb agreement are both present in English and Italian. Yet, both languages differ in the extent they use these devices to mark the subject of the verb. MacWhinney, Bates and Kliegl (1984) studied the effect of word order and morphological marking on the sentence processing by native speakers of German and English and came to the conclusion that to identify the subject of the verb, the German subjects relied more than the English-speakers on morphological cues. This reliance on quantitative variation brings a probabilistic element to the Competition Model: the Competition Model sees language processing as being fundamentally probabilistic in nature (Bates and MacWhinney 1989: 13).

The Competition Model is not a model of competence. In fact, as Bates and MacWhinney (1989: 32) put it, "No functionalist grammar has yet been formulated with enough internal detail to constitute a model of competence in any human language." However, the native speaker's latent knowledge can be described in terms of a network of weighted connections such as correlations between forms and functions and among forms themselves (Bates and MacWhinney 1989: 33). The Competition Model is thus more easily compatible with connectionist views of language.

According to Bates and MacWhinney (1989), there are five key concepts in the Competition Model. First, it assumes that there are two levels of informational structure: a functional level for meanings and intentions, and a formal level for surface forms and expressive devices. Secondly, mapping between form and function is stated as directly as possible. This does not mean that the relationships between form and function are one-to-one; nor does it refer to the direction in which information flows during sentence interpretation. Rather, different sources of information such as phonological or syntactic information can be processed in a similar fashion. The third key concept involves cue validity. Cue validity has three components: availability, reliability (which together determine overall cue validity) and conflict validity. For example, a preverbal position functions as a cue to agent assignment in various languages; in other words, the element that occurs before the verb in a sentence is frequently taken to be the subject of the verb. The validity of the cue depends on how often it is available and how reliable it is, ie. how many times it leads to the correct conclusion and, if there are any competing cues present, how many times this particular cue wins, ie. leads to a correct interpretation. Fourthly, cues have varying strength. Cue strength is the weight on the connections between units. The fifth concept involves coalitions and prototypes. In the Competition Model, it is possible to have many-to-many mappings between forms and functions, which form a series of subsystems which Bates and MacWhinney (1989: 47) refer to as coalitions. For example, the category of sentence subject, or agency, is not to be understood as a unitary concept but a coalition of cues such as animacy, word order and morphological devices, each having other functions as well.

The Competition Model makes a number of claims about language acquisition: First, language acquisition is cue-driven, and the validity of the cues determines the order in which grammatical devices are acquired. For example, the SVO word order in English is the first cue to have a significant effect on sentence interpretation (in children as young as 28 months), whereas semantic cues (such as animacy) are far stronger than word order in Italian (Bates et al. 1984). In other words, children may begin to attend to word order, semantics, and/or grammatical morphology, depending on the relative validity of those cues in their native language. Secondly, conflict validity seems to have an effect on language acquisition: children begin to notice the situations in which two or more cues compete and reset form-function mappings to favor those cues that win such conflict situations. Language acquisition is also affected by changes in 'cue cost'. These are differences in assignability and perceivability of

cues, which may explain delays in language acquisition. Some cues require more memory and processing capacity. For example, Italian children seemed unwilling to use agreement contrasts in sentence interpretation, but relied, instead, on semantic factors. It also appears that certain functions must develop first before mappings from form to function can be made. For example, the conditional and subjunctive forms in Italian are used with their associated meaning quite early during language acquisition; however, the corresponding inflections are not learned until somewhere between 4 and 6 years of age. The cognitive difficulty of the counterfactual notions involved in the use of conditional and subjunctives may explain this result. However, as Bates and MacWhinney (1989) point out, it is dangerous to claim that a given grammatical structure is 'conceptually difficult' without independent evidence.

Much of the research done within the Competition Model has focused on the recognition of the semantic role of the agent in a sentence. In experiments like this, the learner is typically asked to choose the agent of a simple SVO sentence. This method has been used to study both child language acquisition (Bates et al. 1984) and sentence processing by adult bilinguals (McDonald 1987, Kilborn & Cooreman 1987, Harrington 1987, Kilborn 1989). These types of tasks have the advantage of avoiding one of the most typical pitfalls of grammaticality judgment tasks as they involve no error detection. Such tasks can be carried out with or without reaction time measurement. Bates et al. (1984), for example, involved a simple agent selection task by children of various ages.

However, Bates and MacWhinney (1989) point out that more sophisticated methods are needed to investigate the on-line processing of sentences. Even end-of-sentence reaction time data is not enough.

Kilborn (1989) in his study of the sentence processing strategies of Dutch-English bilinguals combined agent selection with reaction time measurement with the expressed purpose of bringing the sentence interpretation task closer to the realm of on-line processing. His results indicated that monolingual English speakers depended on word-order and generally did not attend to morphological information in assigning thematic roles (specifically that of the actor, or agent) during comprehension. In contrast, native German speakers depended on morphological cues in their native language and delayed their responses until all potential cues were in. Furthermore, when the same subjects performed the task in their second language, English, a similar result was observed, indicating that German-based processing strategies were transferred to on-line processing in the second language.

The Competition Model has been criticized for the methodology it uses. The experiments typically involve the use of semigrammatical sentences such as *The horse are hitting the cows* (Bates and MacWhinney 1989), or *The telephones pushes the cowboy* (Kilborn 1989), the agent of which subjects have been asked to decide on. The nature of the nouns and sentences is carefully controlled for; usually, a number of potential cues such as noun animacy, word order and subject-verb agreement is systematically varied in order to get the subjects' responses to an equal number of similar instances of cue occurrence. Also, most of the research done has been in sentence comprehension, not in production. Furthermore, most of the data have come either from judgments of sentences presented on paper or with the help of reaction time measurement at the end of the sentences. For a model in which the timing of cues may be an important factor for the correct calculation of their strength and validity, it is crucial to gather data on-line, during sentence processing. Blackwell et al. (1996) have conducted an exhaustive experiment to investigate the factors involved in on-line L1 grammaticality judgments (see Chapter 7).

2.2. Linguistic theories of second language acquisition: Universal Grammar

Few notions have been so influential in the study of language acquisition as competence and performance. Chomsky first defined these concepts in 1965:

> Linguistic theory is primarily concerned with an ideal speaker-listener, in a completely homogeneous speech community who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of language in actual performance [...] We thus make a fundamental distinction between *competence* (the speaker/hearer's knowledge of the language) and *performance*, the actual use of language in concrete situations. (Chomsky 1965: 3-4)

From the beginning, Chomsky made it clear that the object of linguistics was to study competence, not performance. Later, Chomsky amended his view of competence by introducing the notions of grammatical and pragmatic competence:

> I assume that it is possible to have full grammatical competence and no pragmatic competence, hence no ability to use a language appropriately, though its syntac and semantics are intact. (Chomsky 1980: 5)

> By 'grammatical competence' I mean the cognitive state that encompasses all those aspects of form and meaning and their relation, including underlying structures that enter into that relation, which are properly assigned to the specific subsystem of the human mind that relates representations of form and meaning. (Chomsky 1980: 59)

Chomsky (1986) makes a distinction between externalized language (E-language) and internalized (I-language). While E-language research is interested in the use of language in social contexts and uses such data to describe grammars of languages, the object of I-language research is the study of language knowledge and how it comes about. Moreover, it treats language as an internal property of the human mind (Cook and Newson 1996: 21).

Universal Grammar (UG) model of second language acquisition was based on Chomsky's linguistic theories. Its origins are in the Language Acquisition Device (LAD) proposed by Chomsky (1965) to account for L1 acquisition. In this model, the learner is exposed to language data (input), which goes in to the LAD, the output of which is a fully formed grammar. The UG theory has focused on specifying what the features of the input are, what the processes are within the second language learner, and what the properties of the resultant grammar are (Cook 1993: 200).

Instead of the rules that characterized the early transformational generative grammar, UG describes grammar in terms of principles and parameters. In UG, according to Cook (1993: 201), "the language knowledge of adult human minds takes the form of universal principles and variable settings for parameters, whatever the language they have learnt". Principles are universal and unchanging, no matter what the actual language is. For example, the Projection Principle states that all lexical items project their syntactic specifications on to the phrases of which they are heads. All such principles are part of the language faculty of the mind known as Universal Grammar. However, languages vary. UG explains variation through the existence of parameters. Parameters themselves are universal and can only vary within certain pre-set limits, but they can get different values, depending on the language. Such parameters affect

word order, for example, or whether a language is pro-drop or non-pro-drop, ie. whether a sentence with null-subject is grammatical in a language. In addition to principles and parameters, language knowledge also includes lexical items with their specifications that govern their projection on to the sentence structure.

To account for second language acquisition, there were at first attempts to fit it within the same LAD framework as L1 acquisition. However, second language acquisition differs from L1 acquisition in some crucial ways, the most important being the fact that few second language learners develop a competence equal to that of a native-speaker. Cook (1993: 110) points out that the performance-based views of second language acquisition need to deal with what is knows as cognitive-deficit problem. This term is used to refer to the phenomenon that L2 learners do not perform as well as native speakers on great many linguistic tasks; why this phenomenon is considered a 'problem' may be an indication of how the development of the L2 competence itself is regarded as somehow problematic and ultimately unsatisfactory within the UG framework.

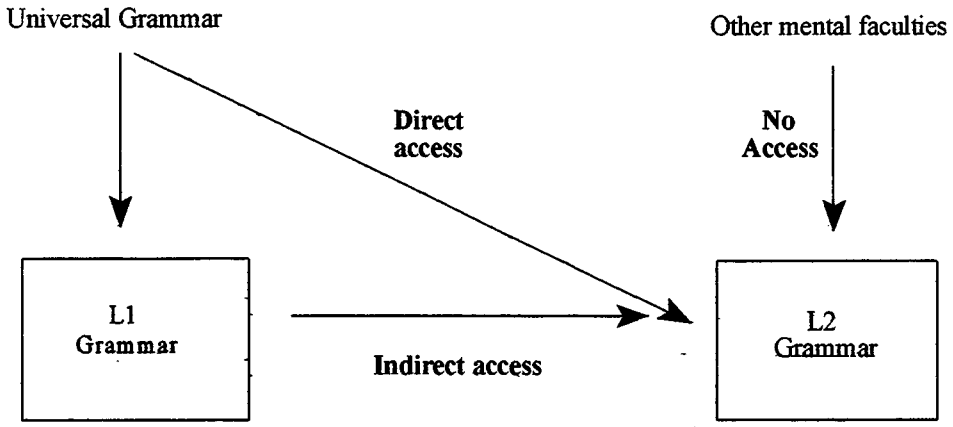


FIGURE 3 Access to UG in theories of second language learning (from Cook 1993: 210).

There are three positions on how UG might work in second language acquisition, described in Figure 3 (for an extensive discussion, see Eubank 1991). The first of these assumes that during second language acquistion, there is a direct access to UG. Second language learners set values for parameters according to the L2 data they are exposed to without any other influence. Hardly anyone holds this position anymore. Secondly, it has been assumed that there is an indirect access to UG. According to this view, the learners have access to UG through their L1 knowledge: learners start with the parameter values set for the L1, which they then try to reset to L2 values. The third position holds that there is no access to UG during second language acquisition (cf. Bley-Vroman, Felix and Ioup 1988). The second language is acquired through other faculties of the mind than language, and consequently, L2 competence is distinct from L1 competence. There are several arguments offered for the no-access view of second language acquisition: first, it explains why L2 knowledge is not complete (Schachter 1988, Bley-Vroman 1989); second, it accounts for fossilization; third, it explains why some languages are more difficult to learn than others; and fourth, it offers a way to account for individual variation in second language acquisition. The explanation why UG is no longer available in second language acquisition is because the

language faculty changes with age (Bley-Vroman 1989). Another hypothesis known as the Critical Period Hypothesis (CPH) is sometimes evoked in connection with this view: according to the CPH, people who acquire a second language before a certain critical age have access to UG and have therefore a different L2 competence than those that acquire a second language after the critical period.

Cook (1993, 1994) has suggested that the confusion in the access research has been caused by a separate learning component that has been implicitly introduced to the UG model.

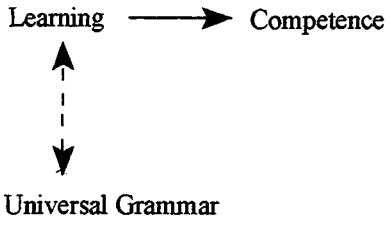


FIGURE 4 Implicit metaphor in access models of L2 learning (from Cook and Newson 1996: 296)

Figure 4 above shows the hidden metaphor underlying most of the research debate on the access to UG.

Cook (1993, 1994) suggests that the whole debate about access to UG is based on false premises: that there exist two separate entities, learning and UG, whereas, as it was originally intended, UG *is* the process through which learning takes place. Second language acquisition research within the UG framework has run into trouble because of the way it has approached the problem; "it asks 'Is there access by some learning process to UG in SLA?' not 'Does L2 learning take place through UG?'" (Cook 1993: 243). Cook, however, also points out that the LAD metaphor is not very helpful either as it rests on the metaphor of the language faculty being transformed into a final state representing one of the possible human languages. But in second language acquisition there is 'no single final steady state', and therefore, no product out of the LAD, only changes in the contents of UG itself (Cook 1993: 243). Rather than debate on issues that may be products of different metaphors, it might be more useful to focus on investigating the acquisition and use of languages and language knowledge.

Cook (1993, 1994) criticizes UG research for its treatment of monolingualism as the norm:

> Why should L2 users be treated as failed monolinguals? Their competence is whatever it is, not to be compared in derogatory fashion with monolingual competence: it is a different kind of thing. (Cook 1993: 244)

and then, later:

> But, if most human beings arguably use more than one language, basing linguistics on monolinguals is not so much a convenient idealisation as a misleading representation of the human species. (Cook 1993: 245)

Cook strongly suggests that the 'multi-competence' of second language learners should be taken into account when the use and acquisition of second languages is studied. For example, the second language learners' target may not be the competence of a native monolingual speaker but instead multi-competence at some level (Cook 1993: 244). This has an important

consequence for principles and parameters: if multicompetence is the norm, it must be possible to express principles and parameters in a form that allows more than one setting, more than one grammar to be held in mind at a time (Cook 1993: 245). If taken to the extremes, it could be even claimed that native speakers have several competencies, for example according to stylistic dimension.

The fundamental difficulty in UG research is that although its object is the study of language learners' competence, its state must be inferred from language learners' performance. Grammaticality judgments have been typically used to gather such data and it has been only relatively recently acknowledged (see eg. Birdsong 1989a, Ellis 1991) that language learners' performance in this type of task is no different than their performance in other types of tasks; the data it offers has no more inherent value than the data obtained by other elicitation techniques.

3 Using second language knowledge

Cognitive approaches to second language acquisition do not make a clear-cut division between the acquisition and use of L2 knowledge. As Ellis (1994: 393) points out, the general assumption has been that the way in which knowledge is stored reflects the way in which it is used. A distinction, however, can be drawn between knowledge as in 'procedural knowledge' and skill as in 'procedural skill' (Schmidt 1992). To account for 'procedural skill' it is necessary to look at how learners use their L2 knowledge and also how they deal with the problems resulting from insufficient L2 knowledge or inability to access L2 knowledge in real time. The role of controlled and automatic processing has been important to much of the discussion in this field (eg. McLaughlin, Rossman and McLeod 1983, McLeod and McLaughlin 1986). Researchers working in this area have also been interested in looking at how learners plan their speech and what type of communication strategies they appear to be using in real-time linguistic tasks. Faerch and Kasper (1986, 1987) have looked at communication strategies used by language learners in different speech situations by using Anderson's concepts of declarative and procedural knowledge. A group of researchers working at the Kassel University (see Möhle and Raupach 1989) have adapted these two notions to theories of second language speech production, as well. Hulstijn (1989, 1990) has applied these notions in second language learning. Much of the discussion in this thesis will evolve around these. Yet another approach to second language acquisition and use is offered by the discussions of second language communicative competence and its relationship to second language proficiency primarily conducted within the framework of language testing. These two very different appoaches can be construed, however, as 'microlevel' and 'macrolevel' discussions of the same phenomenon: second language knowledge and its use. The following discussion will start from the microlevel of language processing, go on to the development of automaticity and how fluency at that level has been construed, to the description of models of second language performance. Finally, issues relating to second language metalinguistic performance will be discussed.

3.1 Language processing

Frazier (1988) distinguishes between three types of language processing theories: *autonomous, interactionist,* and *modular* approaches (cf. Flores d'Arcais 1988: 113-114). The distinction between the first two is related to what Sajavaara and Lehtonen (1980: 28) note on the difference between a static and a dynamic view of language, of which the latter seems to be rather more suitable for the psycholinguistic description and operational modelling of language. In perception, this implies that lexicon, syntax, and phonological structure contribute simultaneously to the process of identification of the speech input, ie. information is processed 'heterarchically', rather than 'hierachically', ie. step by step from a concrete to an abstract level. The knowledge-driven process takes place simultaneously with the input-driven analysis. As Flores d'Arcais (1988: 114) points out, there is a great amount of evidence for top-down effects and for a simultaneous use of sources of knowledge in word and sentence perception. The autonomous model has been linked with Forster (1979). It later proved inadequate in explaining certain empirical results, which

caused such researchers as Marslen-Wilson and Tyler (1980) to develop an interactionist view of language processing. Among the most important properties of interactive models are, for example, that processing proceeds in parallel at all levels, and that the processor is free at any moment to use evidence from any level (Flores d'Arcais 1988: 114). However, some psycholinguists regarded interactive models as too powerful. According to Frazier (1988: 25), they discarded the notions of the serial arrangement of processing subsystems and restrictions on the availability of information prevalent in autonomous models; on the other hand, they retained the claim made in autonomous theories that the language comprehension system consists of several distinct processing subsystems (modules), each with its own properties and its own characteristic information sources, hence the name *modular*. Some views of second language acquisition such as Universal Grammar are strongly in favor of the modular approach. On the whole, as Flores d'Arcais (1988: 114) notes, it is an interesting question whether it is possible to combine the modular view with the existence of interactive processes in perception; the notion of interaction in itself is not incompatible with the idea of specific processing at a given level, at least.

The view adopted in the present study presupposes that the language learner is able to utilize information from all levels of language during the process of visual sentence processing. No claims are made, however, on when and where such information is available for the learner and what it might interact with (cf. Flores d'Arcais 1988: 114). This is not to say that the determination of the exact (if possible) operation of language processing mechanisms is not important, but rather it is merely left undecided as there is no intention of tapping into any on-line subprocesses proper that operate in the sentence processing.

A distinction can be made between 'hierarchical' and 'heterachical' processing; both of these have been described by the terms serial and parallel processing (Fodor et al. 1974). 'Heterarchical' models seem to offer a much more fruitful ground for the study of speech comprehension. The hearer forms hypotheses on the basis of the information he or she is receiving from all levels of linguistic and extralinguistic context. During processing, the hearer is assumed to generate hypotheses about the structure or meaning (interpretations) in two different ways: simultaneously (parallel) and consecutively (serially) (cf. Tommola 1985). Both models have certain consequences concerning the processing load.

It can be assumed that readers/listeners modify their processing in either direction depending on the syntactic structure, semantic content and context of the utterance. Context affects the processing of ambiguous utterances so that it is consecutive (eg. 'garden path' sentences).

Two basic processes are usually held to be involved in sentence comprehension: first, lexical retrieval, and second, parsing, ie. the construction of syntactic representations; to what extent they are two distinct processes is under debate (see eg. Frazier and Rayner 1987; cf. MacDonald, Pearlmutter and Seidenberg 1994). The study of the processes involved in comprehension has proved to be difficult, mainly because they take place very fast on many levels, and largely outside of conscious awareness (Osterhout 1994). In both language comprehension models (eg. Forster 1979, Marslen-Wilson 1987) and speech production models (eg. Levelt 1989) the mental lexicon has a crucial role. In the following, the structure of the mental lexicon is discussed in the light of the speech production model described by Levelt (1989). Figure 5 shows the components of Levelt's model.

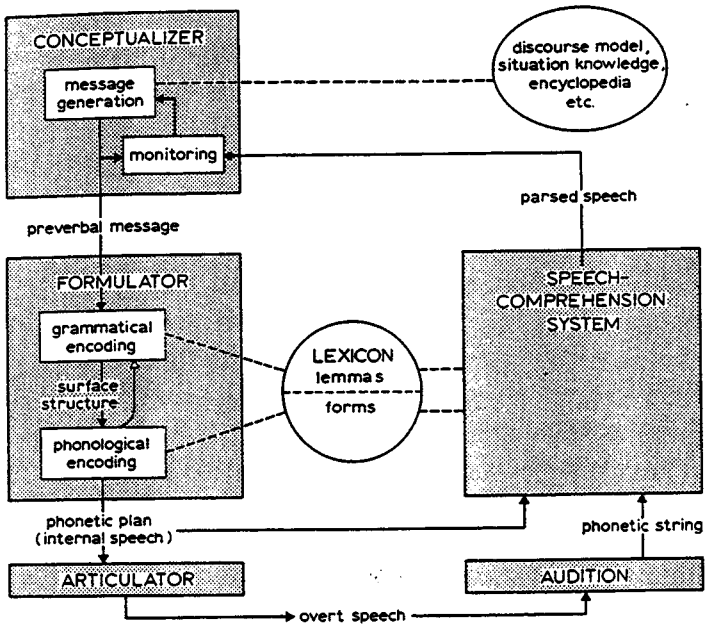


FIGURE 5 Levelt's blueprint for the speaker (from Levelt 1989: 9).

According to Levelt, the speaker has two kinds of knowledge: procedural and declarative. Speakers' lemma information is declarative knowledge, contained in their mental lexicon (Levelt 1989: 11). A lexical item's lemma information consists of the lexical item's meaning or sense and its syntax (for example, the verb *give* is categorized as a verb which can take a subject expressing the actor, a direct object and an indirect object expressing the recipient). A lexical item also contains information about its internal composition, its lexical form. This information is morphological and phonological. The knowledge used in grammatical encoding is procedural; it is stored in the Grammatical Encoder. Basically, it contains IF-THEN instructions on how to build verb and noun phrases on the basis of the syntactic information. Similarly, Phonological Encoder contains procedural knowledge primarily used to access information about the lexical form (Levelt 1989: 12). In language comprehension, parsing is a process that is often viewed as operating on the purely processual level of language, separate from linguistic knowledge (White 1991: 169). Yet, the language user's L1 knowledge has been shown to have an indirect influence on the assignment of L2 syntactic structure (see eg. Bates and MacWhinney 1989)

The exact relationship between language learning and language processing is not quite clear although it is well-known that it is easier to study language processing than processes involved in language learning. Furthermore, as Genesee (1988: 93) points out,

> Although it is reasonable to suspect that the same neuropsychological substrates underlie both language learning and processing, this is not necessarily the case [...] the neuropsychological processes involved in language learning may not be the same as those involved in using a second language for normal communicative or intellectual purposes. Whether or not they are the same is an empirical question. At present there appears to be no way to examine the neuropsychological bases of language learning themselves; at best, researchers can investigate language processing. (Genesee 1988: 93).

In other words, the study of mental learning processes themselves is almost impossible. Attempts have been made recently, however (see eg. N. Ellis 1994, N. Ellis and Schmidt 1997).

3.2 Development of automaticity

Automaticity was originally regarded as a dichotomy of automatic and controlled processes (Posner and Snyder 1975, Shiffrin and Schneider 1977). Since then, it has become more usual to view it as a continuum. Such phenomena as speed-up in reaction time and error rates and diminished interference from simultaneous tasks are held to be reliable criteria of automaticity (DeKeyser 1996: 350). (Not everyone agrees: Segalowitz and Segalowitz (1993), for example, make a distinction between speed-up of strategic processes and automaticity.) The term automatization refers to the gradual development of automaticity through practice. An important issue for second language acquisition is whether automatization applies only to items or whether the application of abstract rules, or algorithms, can also become more automatized (DeKeyser 1996). Below, some of the major views on the development of automaticity will be dicussed.

First, it should be noted that it is important to make a distinction between automaticity and skill (Logan 1985). Skill is a term usually applied to performance of a complex task; the task itself is regarded as a skill, and those who perform better on it, are considered more skilled than those who perform worse. On the other hand, automaticity refers to specific aspects of performance. Usually, tasks that can be performed quickly, effortlessly, and relatively autonomously are thought to be automatic. Skilled performance utilizes a large number of automatic processes. However, skills do not consist of automatic procedures alone but also of metacognitive knowledge about how and when to use the procedures (Logan 1985: 369).

3.2.1 Controlled and automatic processing

The notion of controlled and automatic processes has proved to be very important in the field of second language learning. As mentioned above, if second language learning is regarded as the acquisition of a complex cognitive skill, then, like any other complex cognitive skill, it can be considered to involve the gradual integration of subskills, which are slowly automatized in order for processing capacity to be freed for other mental operations.

According to Shiffrin and Schneider (1977), there are two modes of human information processing, controlled and automatic, each operating on the nodes in memory. Human memory is viewed as a collection of nodes that are sets of informational elements. These nodes become associated through learning. The nodes can be inactive and passive, in which case this network of nodes is called the long-term store. If, because of some kind of external stimulus, a number of these nodes are activated, the activated nodes form the short-term store. Automatic processing involves the activation of certain nodes in memory every time the particular inputs are present; the activation is a learned response that has required a certain amount of training to develop. There is a qualitative difference between the two modes of processing: automatic processing is fast, effortless, autonomous, and not available to conscious awareness (Logan 1988: 492). According to Shiffrin and Schneider (1977: 127), automatic processes can be characterized as demanding little or no attention, parallel in nature, difficult to alter, ignore or suppress once learned, and unaffected by processing load. Controlled processing, on the other hand, is a temporary activation of nodes in a sequence. Controlled processes demand a great amount of attentional capacity, are usually serial in nature, easily established, altered, and strongly dependent on load.

As McLaughlin (1987: 135) points out, controlled processes seem to regulate learning, which involves the transfer of information to long-term memory. Any skills are learned and routinized, that is, become automatic, only after the earlier use of controlled processes. Once such automatic processes have been created, controlled processes can be allocated to higher levels of processing. Speaking is a good example of the interaction between automatic and controlled processing (cf. Levelt 1989: 20-21).

McLeod and McLaughlin (1986) have applied the notion of automatic and controlled processing in the study of L2 reading. Their framework differs from the traditional view of automaticity developed by Shiffrin and Schneider (1977) in that the restructuring of knowledge and consequently also of the processes used to deal with that knowledge is an essential part of it.

Automaticity is linked to the study of attention. Attention can be viewed as an aspect of a single-capacity system as in Kahneman's (1973) single-capacity model, or it may be regarded as a process or ability that can be used more or less efficiently (Just and Carpenter 1992). The exact meaning of 'paying attention' is not clear, however. As Blanco and Alvarez (1994: 78) point out, a difference has to be made between several varieties of attention such as sustained and selective attention; in addition, attention may be partially dependent on modality, ie., visual and verbal stimuli are attended to in a different way. Within the single-capacity framework, automatic processes are regarded as occurring without attention; in fact, the acquisition of automaticity is the gradual withdrawal of attention (eg. Shiffrin and Schneider 1977). Automatic processes do not require capacity, which attention implies, which is why they are fast and unavailable to control, and since attention is the mechanism of consciousness, they are unavailable to consciousness, as well (Logan 1988: 492-3). This notion of automaticity has since been challenged, either by questioning the single-capacity model as the multiple-resource theories do, or attacking the notion that automaticity implies withdrawal of attention (Logan 1988: 493). Moreover, not all researchers make a distinction between automatic and controlled modes of information processing. According to the view usually held, automatic processes are difficult to control, and when the decisions or responses required by the task conflict with such automatic processes, the latter interfere with performance by typically increasing reaction time. However, Logan (1985: 378-379) points out that it seems that automatic processes can be controlled as skilled performers are usually able to control their performance better than unskilled performers, even though their performance still retains features typical of automatic processes. For example, adult speakers have very close control over their speaking although such a highly skilled activity as speaking is best characterised as automatic. More generally, the language learner's processing is largely automatic when her language proficiency is good. Thus, according to Logan (1985: 367), automaticity and skill are not to be identified with each other.

3.2.2 Instance theory

To offer an alternative explanation to automaticity, Logan (1988; see also Logan and Klapp 1991, Logan and Etherton 1994) has developed a theory of automaticity known as instance theory. The instance theory of automatization is based on the idea of automaticity as retrieval from memory. By viewing automaticity as part of memory system, the theory is able to avoid the implications inherent in the view of human information processing as the single-capacity system with limited resources:

> Automaticity is memory retrieval: Performance is automatic when it is based on single-step direct-access retrieval of past solutions from memory. [...] novices begin with a general algorithm that is sufficient to perform a task. As they gain experience, they learn specific solutions to specific problems, which they retrieve when they encounter the same problem again. Then, they can respond with the solution retrieved from memory or the one computed by the algorithm. At some point, they may gain enough experience to respond with a solution from memory on every trial and abandon the algorithm entirely. At that point, their performance is automatic. (Logan 1988: 493).

Logan (1988) gives children's acquisition of simple arithmetic as an example: at first, children learn to add single-digit numbers by counting, one by one. Gradually, children learn the sums of all pairs of single-digit numbers by rote. After that, they no longer need to count to get the correct answer; instead, they rely on their memory.

According to Logan (1988), instance theory assumes that attention has two unavoidable consequences: first, encoding into memory, and second, retrieval from memory. In other words, attending to a stimulus is sufficient to commit it to memory (cf. Schmidt 1990, 1994); likewise, attending to a stimulus is sufficient to retrieve from memory whatever has been associated with it in the past. This does not mean that everything that has been attended to is encoded as it is: attention can be of differing quality, and the quality of encoding depends on the the quality and quantity of attention; thus, not everything is encoded equally well. Nor does it mean that everything that has been attended to is remembered: retrieval is not always successful, either.

The theory also assumes that every time a stimulus is encountered, it is also encoded, stored and retrieved separately. Learning takes place through the accumulation of such memory traces, which is assumed to gradually lead from algorithmic processing to memory-based processing; hence the term *instance* theory. The memory traces are linked to specific stimuli: consequently, in instance theory, automatization is item-based. Transfer to new stimuli and situations is difficult, whereas there are no such limitations in the process-based views of automatization such as Shiffrin and Schneider's (1977). As Schmidt (1992: 371) points out, it is a major weakness of instance theory that it cannot explain transfer of training or generalization in skill learning.

Logan (1988: 517) points out that according to his view, it its likely that "humans can learn in more than one way". Thus, the instance theory of automatization may better explain the types of learning where only one step, or procedure is involved. Other theories such as Anderson's ACT* model are better suited to describe learning processes that involve the transition form multi-step procedures to simpler procedures (as in Anderson's composition stage), or situations in which people learn general procedures rather than specific responses to specific stimuli (Logan 1988: 516). Logan (1988) indicates quite clearly that in his view, both memory-based and algorithm-based processes are not mutually exclusive but exist side by side, and are each used depending on the nature of the problem to be solved.

Lexical decision tasks are a good example of a task suitable for automatization in the way instance theory proposes. However, higher-level units than words may be retrieved from memory such as phrases and both completely formulaic and partly open clausal strucutures (Schmidt 1992: 371). Schmidt (1992), in discussing fluency in speech production, suggests that instance theory could provide theoretical support for Pawley and Syder (1983: 218), in whose view nativelike fluency in speech is based on the memorization of hundreds of thousands of complex lexical items, including lexicalized sentence stems; for example, holistically stored sequences have the "advantages of being quickly retrievable and

of being familiar to the hearer as well as the speaker". Likewise, it also supports Widdowson's (1989) suggestion that second language learners' use of linguistic knowledge for communication may rely more on formulaic frameworks or partially preassembled patterns retrieved from memory rather than on using rules to compose utterances step by step.

Robinson and Ha (1993) applied Logan's instance theory to adult second language acquisition. The experiment was conducted by using a special program design to measure subjects' reaction times and grammaticality judgments. In the experiment, they first trained a group of English L2 learners in the dative alternation based on the morphological explanation initially devised by Mazurkewich and White (1984). According to the explanation (Robinson and Ha 1993: 418), verbs with only one syllable (eg. *give*) in their base form can occur in two types of sentences: first, in sentences like *John gave Mary a cake*, and second, in sentences like *John gave a cake to Mary*. Verbs with two syllables (eg. *donate*) however, can only occur in the latter type of sentence. The rule was given twice, with examples. After that the subjects were shown sentences with made-up verbs one at a time, the grammaticality of which they were asked to judge by pressing a key on the computer keyboard. For example, the subjects were asked to judge the grammaticality of sentences like *John nawked Mary the ball*, *Bill reliped the letter to Jack* and *George menided Harry the pen*, of which the last one was ungrammatical according to the rule. After each instance, the subjects were given feedback on the screen as to whether their response had been correct. Artificial and nonsense verbs were used to prevent any previous exposure to real English verbs from interfering with the results. The type and frequency of the sentences were also carefully controlled for. In the second part of the experiment, the aim was to see whether the learners would show the effects of frequency and transfer to new instances in their responses to sentences containing old and new verbs and frames of the type they had been trained to judge. Again, the type and frequency of the verbs and also the frames surrounding these verbs were carefully controlled for.

Instance theory does not predict any accuracy difference, but it predicts a difference in the speed of the response (Robinson and Ha 1993: 417). The subjects' responses failed to show the effect of frequency, ie. the verbs that had been showed fewer times did not receive higher RTs than those that had been presented more frequently. However, the subjects' judgments of old instances were significantly faster than their judgments of the new ones. In other words, the accumulation of memory traces predicted by Logan's theory manifested itself in the subjects' responses. There was also some evidence of verbs and their frames that the familiarity with the frame had a speeding effect on the response times; it was not as great as the familiarity with the whole instance but greater than that with the verb. There was no difference in the speed with which the learners judged monosyllabic and disyllabic verbs, nor were significant differences found between the processing speed of frames with monosyllabic and disyllabic verbs. Robinson and Ha (1993: 426) interpreted the latter results as evidence against the verb-first strategy in parsing, and suggest that attention is distributed equally over the whole frame during the processing. More importantly, though, for the purpose of the experiment, the significant differences in the speed of judgments on old and new items implies that there is no transfer of automaticity: automaticity developed during the training session on old instances did not transfer to new instances enountered during the training phase.

Robinson and Ha found no statistically significant differences in the accuracy of judgment: monosyllabic vs. disyllabic verbs, old vs. new instances and the changing of verb frames appeared to have no significant affect on the subjects' accuracy. Robinson and Ha

(1993: 427) interpreted these results as evidence for the transfer of the algorithm developed during the training phase to new instances in the transfer phase. Taking these two results from RT and accuracy data, Robinson and Ha concluded that these two processes appear to work hand in hand in the decision-making process such as the grammaticality judgments. It appears that these two processes work parallel, as Logan (1988) suggested. Robinson and Ha (1993: 432) conclude by pointing out that "a transition from algorithm-based processing to memory-based processing occurs and is responsible for the speedup in skilled performance on grammaticality judgments about the examples presented". Second language learners can have explicit rules that they apply in the decision-making process, but the greater speed of decisions is ultimately due to retrieval from memory.

Robinson and Ha (1993: 433) raise two questions. The first involves the targets and timing of automatization: which language structures are subject to algorithm-based and memory-based processing and when does the transition to memory-based processing take place? It is possible that the existence of these two systems could shed new light on some of the findings on the developmental sequences in acquisition (see eg. Meisel, Clahsen and Pienemann 1981, Pienemann, Johnston and Brindley 1988, Pienemann 1989).

Second, their findings have certain methodological consequences for studies using reaction time measurement to complement grammaticality judgment data to find evidence for the existence of two competing rules in the learners' interlanguage system (which the experiments in this thesis were not used for). The decisions made by subjects in such studies, typically conducted within the UG framework, could be memory-based as well as algorithm-based. What the implications would be for the Competition Model is unclear; in the experiments set up by researchers working within the model, the subjects are often asked to make decisions concerning the agent of the sentence. If automaticity is based on the retrieval of instances stored formulaically in memory, fast decisions of a certain kind on this type of task could be a reflection of more accumulated instances of a certain type in memory and not a function of animacy or subject-verb agreement (ie. morphological devices) as such. The Competition Model, on the other hand, makes no claims about the existence of rules, or even algorithms, in the knowledge system.

3.2.3 Connectionism and the Competition Model

A major distinction can be made between instance theories and strength theories in the human information processing. Logan's instance theory is an example of instance theory developed to explain automatization (there are others, for example for categorization). Connectionist models like the ones developed by McClelland and Rumelhart (1985) and Cohen, Dunbar and McClelland (1990) are strength theories. Other strength theories have been proposed by Schneider (Schneider 1985; Schneider and Detweiler 1988). The difference between instance and strength theories lies in the way acquisition is considered: instance theories regard it as an accumulation of memory representations, or instances, while strength theories regard it as the strengthening of connections between stimulus and response in the neural networks during practice. The notion of distributed or shared representation is central to many connectionist theories: according to McClelland and Rumelhart (1985) and Cohen, Dunbar and McClelland (1990), memory for events (ie., episodic trace) is not encoded as a discrete instance, not even as a single connection between a stimulus and a response, but in the strengths of connections between different units; these

connections are then used to provide distinct yet overlapping representations. Cohen, Dunbar and McClelland (1990) have proposed that the role of attention in task performance is to select among competing processes on the basis of task instructions; thus, attention serves as an additional source of input that affects interactions occurring at the intersections of connectionist pathways.

Following Bates and MacWhinney's (1989: 51-59) account, the Competition Model assumes that the mapping of form to function during sentence comprehension is governed by a system of parallel activation in which conflicts are decided on the basis of strength. The resolution of such competing processes accounts for the assignment of grammatical roles and other decisions. For example, when a listener or reader is presented with sentences with varying word order and asked to assign (ie., map) the role of the agent to a word in that sentence, all word order variations are activated parallel to each other. However, English speakers rely much more on word order than Italian speakers, who rely more on semantic contrasts and subject-verb agreement. The mappings are constantly updated, as well. For example, at each point in the processing of the sentence *The dogs are chasing the cat*, the mapping from the lexical item *dogs* to the agent role is updated as different cues (word order, subject-verb agreement etc.) are processed; in this case, the strength of the agent role assignment increases by each updating. It seems that in sentence comprehension, the listener does not simply add up cues in a bottom-up fashion, though: for example, the reaction times can be slowed by cues that the listener typically does not use, like subject-verb agreement in English.

There are limitations to processing, the most important of which are perceivability and assignability. It seems that the degree to which the listener is able to perceive the cue interacts with overall cue validity in determining the probability of a particular interpretation. Schmidt (1990, 1994), in his discussion of noticing, makes a similar point of the importance of saliency for attention and learning. Furthermore, memory load increases during sentence processing when the integration of available material must be delayed until more information is received. The amount of material that must be held in memory before a meaning assignment can be made affects the processing of the cue. For example, Italian speakers cannot make use of number agreement in the third person until they have heard the verb and all its associated nouns; it may turn out that the agreement cue provides no information at all. Thus, cues with low assignability may become so costly that they are abandoned despite their information value.

Bates and MacWhinney (1989: 33-34) discuss the question if a connectionist model can be combined with traditional rule-based accounts of cognition. They present the arguments by Pinker and Prince (1988), who describe three possible relationships: First, connectionism could be implementational in nature: connectionist representations would only show how symbols and rules might be implemented in a neural network. Secondly, connectionism could eliminate the need for higher-order decsriptions such as 'symbols' or 'rules', which would have value only as descriptive tools but no explanatory value. Thirdly, there could be some sort of symbolic-revisionist connectionism. This comes closest to Bates and MacWhinney's view. According to their view, both levels of description are important. The underlying facts of human performance (subsymbolic level) constrain the set of competence models (symbolic theories).

3.3 Fluency

Fluency can be defined in various ways (see Koponen 1992 for an excellent discussion). Bachman and Palmer (1984: 35) see speech rate and fluency as a component of language skills concerning the speed and ease with which the language user can employ language knowledge. The term fluency is often used to refer to the ability of non-native speakers to speak a foreign language. Schmidt (1992: 358) makes a difference between such "global conception of fluency, which appears to differ little from the concept of proficiency" and fluency as "automatic procedural skill". This term, which Schmidt prefers to use, comes from Carlson, Sullivan and Schneider (1989), who have studied the effect of practice and working memory in skill learning.

Segalowitz (1986, 1991) has discussed the issues related to fluency in receptive processing. Segalowitz (1986), in discussing L2 reading difficulty, suggests that certain basic cognitive mechanisms underlying reading function less efficiently during the processing of L2 material in contrast to the processing of L1 material. Segalowitz gives an indirect definition of fluency:

> By "fluent bilingual" is meant the person who has for all practical purposes rapid and accurate ability to use the vocabulary and syntax of a second language, at least when required to perform under normal speaking and listening conditions, and is also generally skilled at reading the second language [...] people who can express most ideas equally well in each language, who demonstrate good mechanical fluency in the second language although they may possess a slight accent [...] and whose reading rate when reading for general comprehension is within the normal range [...] (Segalowitz 1986: 4)

Thus it appears that according to Segalowitz, there are two dimensions involved in fluency: speed and accuracy. Lehtonen, Sajavaara and May (1976) make a similar distinction, as well. There are other definitions, though, according to which fluency in speech refers to the speed of processing only (Schmidt 1992).

Segalowitz (1986: 7) relates the speed to the development of automaticity after extensive practice. The notion of accuracy appears to him be inherent in skill as "one important distinction between skilled and less skilled readers is the availability of fast automatic word recognition skills in the former" (Segalowitz 1986: 7). Likewise, Segalowitz (1991) in his study of the effect of advanced L2 skill on L1 automaticity makes a distinction between a given skill, such as visual word recognition, and the psychological processes underlying it. Referring to Shiffrin and Schneider (1977) and Schneider and Dumais (1981), he makes a similar distinction between automatic and controlled processes (he prefers to use the term strategic processing for the latter term). Segalowitz (1991: 63), however, points out a difficulty in using fast response as the primary indicator of automaticity since, according to him, practice may very well reduce response latencies in nonautomatic, strategic processing as well. Thus, it appears that he views the two processing modes as independent, like Posner and Snyder (1975), while Schneider and Detweiler (1988), for example, view automatization as a gradual transition from more controlled processing to fully automatic processing.

In his discussion of fluency in speech production, Schmidt makes the following points on fluency: first, fluency depends on procedural knowledge (Faerch and Kasper 1984); and second, if a distinction is made between procedural knowledge (the knowledge of how something is to be done) and procedural skill ("the performance aspect of actually doing

something in real time"), fluency should rather be identified with procedural skill (Schmidt 1992: 359). According to Schmidt (1994: 21),

> Spontaneous, fluent language performance is unconscious only in the sense that it is accomplished without the conscious retrieval of explicit knowledge that may have been used as an aid to production in earlier, novice stages of development. Automatic use is related to both the familiar notion of procedural knowledge and the psychological construct of "implicit memory" [...] but is not by itself evidence for implicit learning.

Thus, it appears that no claims can be made as to the nature of the learning process itself on the basis of observed automatic performance.

3.4 Models of second language performance

According to Bachman (1990: 16), the term language proficiency has been traditionally used in the context of language testing to refer to "knowledge, competence, or ability in the use of a language, irrespective of how, where, or under what conditions it has been acquired". Communicative competence is a broader term; recently, the term proficiency has come to be associated, at least in the foreign language teaching context, "almost exclusively with the a specific language testing procedure, the ACTFL/ILR Oral Proficiency Interview".

Writer	Model of knowledge	Model of performance	Actual use
Hymes (1972)	communicative competence		performance
	knowledge	ability for use	
Canale and Swain (1980)	communicative comptence	[unable to be modelled]	communicative
Canale (1983a, b)	communicative competence		actual communication
	knowledge	skill	
Bachman (1990)	communicative language ability		
	(language competence/ knowledge)	(strategic competence)	

FIGURE 6 Major models of second language performance and their relationship to Hymes' s communicative competence (from McNamara 1996: 60).

Because of the danger of terminological confusion, the term that Bachman prefers to use is language ability (1990: 16). In the following, some of the basic concepts pertaining to language competence and performance will be discussed.

The distinction of competence and performance first set up by Chomsky in 1965 has since been revised and refined by researchers working in many different fields. In 1972, Hymes gave a new formulation to these notions bringing about a theory of communicative competence, a concept that has since come to dominate discussions on the nature of second language performance. McNamara (1996) has reviewed the often confusing and vague use of the terms and notions involved in the discussions of the relationship between competence, performance and proficiency.

According to McNamara (1996: 48), in all these discussions there are basically three dimensions involved: First, there are the factors that constitute *knowledge* of a language. Second, there are the underlying factors forming part of an individual's ability to perform communicative tasks which Hymes (1972) called *ability for use*. Third, there is the relationship between *real-time instances of language use* and these two groups of factors.

McNamara (1996: 60) presents three influential models of second language performance and their relationship to Hymes's (1972) theory of communicative competence. Figure 6 above shows the major dimensions of each model in this classification, with some modifications.

3.4.1 Canale and Swain's model

Hymes (1972) developed his theory to explain native speakers' communicative competence. Canale and Swain (1980) were the first to extend his theory to a comprehensive discussion of the nature of second language communicative competence. According to Canale and Swain, communicative competence is limited to the knowledge of language. It consists of three components:

1. grammatical competence, which includes "knowledge of lexical items and of rules of moprhology, syntax, sentence-grammar semantics, and phonology" (Canale and Swain 1980: 29).
2. sociolinguistic competence, which includes knowledge of sociocultural rules of use.
3. strategic competence, which includes various "'coping' strategies" (Canale and Swain 1980: 31) to be used in actual performance if any of the other areas turn out to be inadequate. Canale and Swain (1980: 7) quite explicitly state that they have excluded ability for use from their model since, in their view, it cannot be modelled. There is no potential for use or underlying skill included in their model. Communicative performance is defined as the realization of the three subcompetences and their interaction in the production and comprehension of utterances, "the actual demonstration of this knowledge in *real* second language situations and for *authentic* communication purposes" (Canale and Swain 1980: 6). As McNamara (1996: 62) points out, it seems that what Canale and Swain mean by communicative performance is simply actual use. Also, 'coping' strategies appear to be nothing else than ability or skill. Furthermore, how the two competences interact is left unspecified, as Canale (1983a: 12) acknowledges.

Canale (1983a: 6) amended the earlier model by distinguishing actual communication, as communicative performance is now called, from the knowledge and skills underlying it:

> Both knowledge and skill underlie actual communication in a systematic and necessary way, and are thus included in communicative competence [...]This view is not only consistent with

> the distinction between communicative competence and actual communication but depends crucially on it; in particular, this notion of skill - how well one can perform knowledge in actual situations - requires a distinction between underlying capacities (competence) and their manifestation in concrete situations (actual communication). (Canale 1983a: 6)

Thus, Canale (1983a) now uses the term communicative competence for both Hymes's notions of language knowledge and ability for use. Canale (1983a 1983b) also introduces discourse competence as a fourth subcomponent of language knowledge, which he defines as

> mastery of how to combine grammatical forms and meanings to achieve a unified spoken or written text in different genres. (Canale 1983a: 9)

However, strategic competence remains rather vague in nature; as a matter of fact, they now seem to include communicative strategies:

> Of course such strategies need not be limited to resolving grammatical problems: actual communication will also require learners to handle problems of sociolinguistic nature [...] and of discourse nature (e.g. how to achieve coherence in a text when unsure of cohesion devices). (Canale 1983a: 11)

Canale and Swain (1980) and Canale (1983a, 1983b) make no mention of the declarative and procedural knowledge. Yet, it appears that some of the contradictions apparent in their theory might have been solved by making a distinction between procedural knowledge and procedural skill. Still, despite its weaknesses, Canale and Swain's (1980) model was the first of its kind to attempt a full description of second language learners' language knowledge and use. It serves both as a basis on which subsequent reseachers have built their models, and as a against which many of the newer models are judged.

3.4.2 Bachman's 1990 model

Bachman (1990) proposes a model of communicative language ability, primarily to satisfy the need for a theoretical framework incorporating the notion of language proficiency with the methods ofr testing it in the context of language testing. It not only recognizes both knowledge of, or competence in the language and the capacity for using this competence but also attempts to "characterize the processes by which the various components interact with each other and with the context in which language use occurs" (Bachman 1990: 81).

The outline of Bachman's model is shown in Figures 7 below. There are three components in Bachman's model of communicative language ability:

1. Language competence.
2. Strategic competence
3. Psychophysiological mechanisms/skills

Bachman's model separates knowledge of language from general cognitive skills involved in language use, for which he uses the term strategic competence. Bachman (1990: 106) defines strategic competence as

> a general ability, which enables an individual to make the most effective use of available abilities in carrying out a given task, whether that task be related to communicative language use or to non-verbal tasks such as creating a musical composition, painting, or solving mathematical equations. Bachman (1990: 106)

Bachman bases his model of strategic competence on the psycholinguistic model of speech production by Faerch and Kasper (1983); he extends it, however, from interlanguage communicative situations to more general communication (Bachman 1990: 100). Faerch and Kasper (1983) based their model on the work of Miller, Galanter and Pribram (1960) and Clark and Clark (1977). Faerch and Kasper's model included two major phases, a planning phase and an execution phase. Bachman's strategic competence consists of three major components: assessment, planning, and execution. According to Bachman (1990: 100-1), the assessment component enables language users to identify the information needed for realizing a particular communicative goal, determine what language competencies are available for the most effective realization of that goal, ascertain the abilities and knowledge shared by interlocutors and, after the actual utterance, evaluate the extent to which the communicative goal has been achieved. The last function resembles Krashen's monitor or psycholinguistic monitoring, but Bachman's assessment is a vastly broader concept. Similarly, his definition of psychophysiological mechanisms involved in communicative

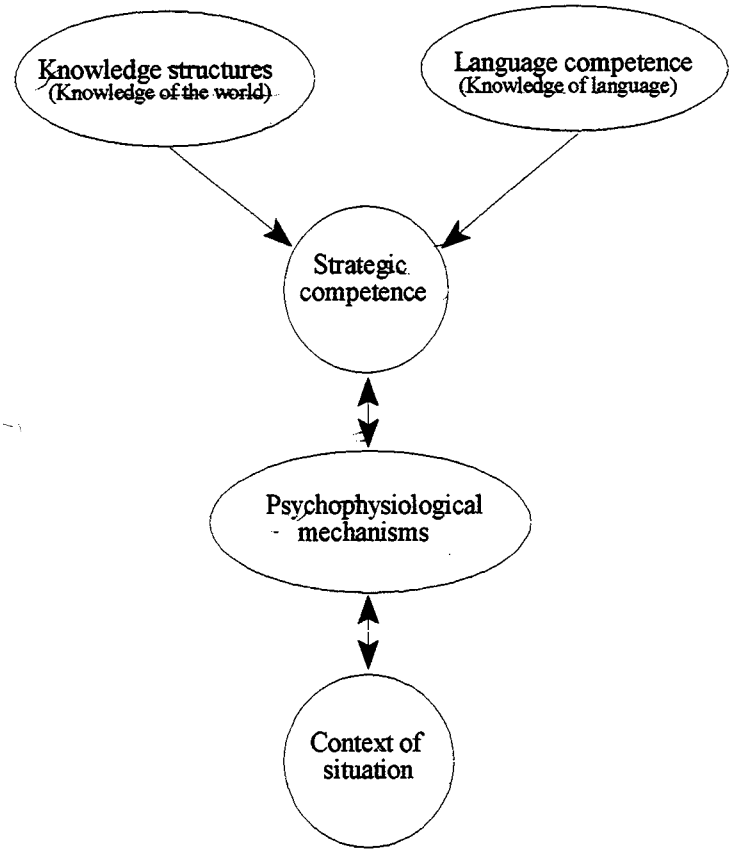


FIGURE 7 Components of communicative language ability in communicative language use (from Bachman 1990: 85).

language ability is based on Faerch and Kasper's (1983) discussion of the executive phase of language use. It includes neurological and physiological processes, and distinguishes the visual from the auditory channel and the productive from the receptive mode. For example, in "receptive language use, auditory and visual skills are employed, while in productive use the neuromuscular skills (for example, articulatory and digital) are employed" (Bachman 1990: 107).

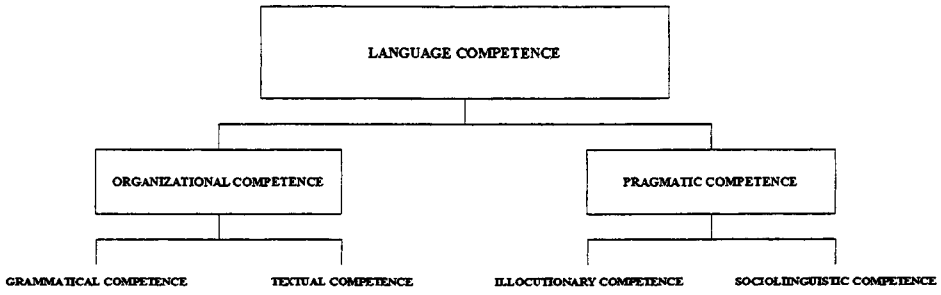


FIGURE 8 A simplified model of language components (based on Bachman 1990: 86).

Bachman's model of language competence is very detailed. Figure 8 shows the main subcomponents of language knowledge, with the last level of subcomponents excluded. Grammatical competence, the area closest to the present research interest, consists of a

> number of relatively independent competencies such as the knowledge of vocabulary, morphology, syntax, and phonology/graphology. These govern the choice of words to express specific significations, their forms, their arrangement in utterances to express propositions, and their physical realizations, either as sounds or as written symbols. (Bachman 1990: 87)

Bachman (1990: 87) gives an example how this might work in language production during a communicative task: the speaker (Bachman uses the term test taker), when asked to describe the relative difference in size between a boy and a girl in a picture, demonstrates her lexical competence by choosing words with appropriate meanings (*boy, tall, girl*), her morphological knowledge by affixing the inflectional morpheme *-er* to *tall*, and her syntactic knowledge by putting the words in the proper order, to compose the sentence *The girl is taller than the boy*. She uses "the phonological rules of English" to produce an utterance that accurately represents the information in the pricture (Bachman 1990: 88). There are two things to be noted about Bachman's model of language competence: first, he appears to be mainly concerned with language production; and secondly, he seems to view language knowledge as consisting of relatively separate modules of vocabulary, morphology, syntax and phonology. In contrast, for example, Levelt's (1989) model of speech production considers all lemma and form information to be contained within lexical entries in the mental lexicon.

3.4.3 Bachman and Palmer's 1996 model

Bachman and Palmer (1996) have revised Bachman's 1990 model to include for the first time, within an explicit model of language use, the role of affective or volitional factors in language use (McNamara 1996: 74). Figure 9 shows the relationship of language use and language test performance.

The former knowledge structures are now labeled topical knowledge. There are relatively minor changes in the way language knowledge is viewed. Bachman (1991) started to use the term knowledge instead of competence some years earlier. The illocutionary subcompetence of the 1990 model has now been replaced by functional knowledge. There have been major changes, however, in the way ability for use is modelled:

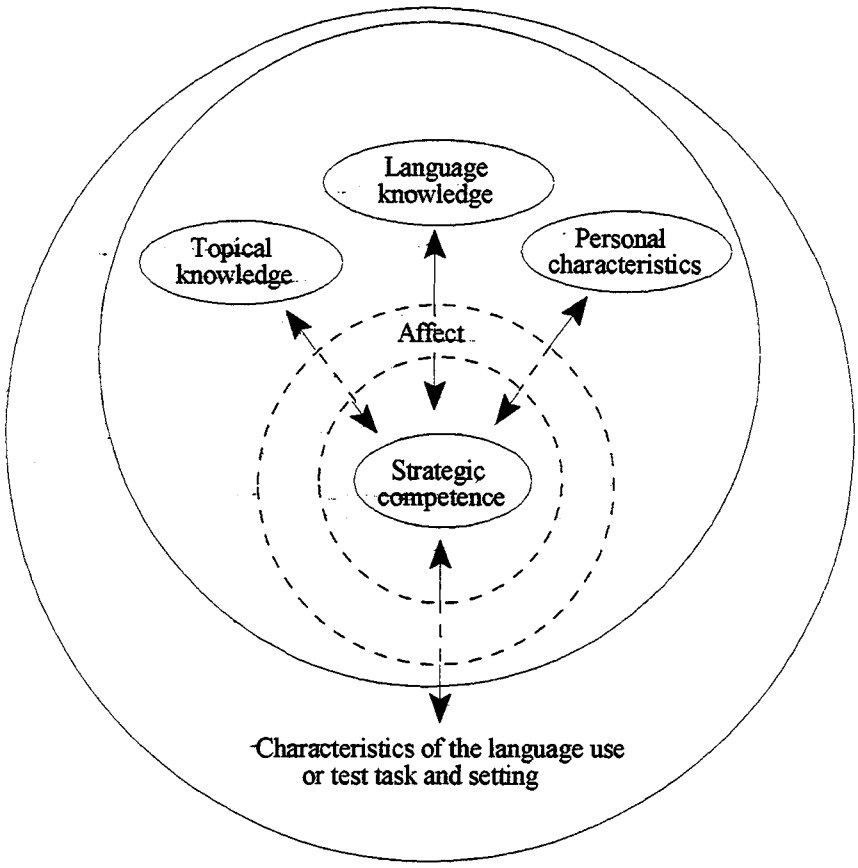


FIGURE 9 Some components of language use and language test performance (from Bachman and Palmer 1996: 63)

A new component called affect, or affective schemata, has been added. They have been referred to as "the means by which language users assess the characteristics of the language use task and its environment in terms of past emotional experiences in similar contexts." (Bachman and Palmer 1996: 65). The function of affective schemata is described as follows:

> The affective schemata, in combination with the characteristics of the particular task, determine, to a large extent, the language user's affective response to the task, and can both facilitate and limit the flexibility with which he responds in a given context. The affective responses of

> language users will thus determine not only whether they even attempt to use language in a given situation, but also how flexible they are in adapting their language use to variations in the environment. (Bachman and Palmer 1996: 65)

In other words, there are now cognitive and non-cognitive aspects included in the ability for use. McNamara (1996: 74) classifies Bachman and Palmer's new dimensions within ability for use in two categories: those that involve knowledge, ie. topical knowledge and affective schemata, both of which are "cognitive structures (essentially memories) and in this respect resemble *language knowledge*"; and those that are processes, ie. strategic competence, which consists of areas of metacognitive strategy use. The metacognitive strategies recognized in the 1996 model include goal-setting, assessment, and planning. The executive aspect has gone, as well as psychophysiological mechanisms primarily connected with it. As McNamara (1996: 74) points out, the significance of affective schemata is discussed mainly from a practical viewpoint as to what its consequences are for the validity of language testing. Also, in some ways the function of affective schemata is viewed in a way similar to Krashen's affective filter (McNamara 1996: 75).

3.5 A psycholinguistic model of second language performance

The previous chapters have described different aspects of second language performance from the level of language processing to communicative language use in test performance. Models of speech production and comprehension and theories of automatization and skilled performance focus on the description of the mechanisms operating at the microlevel of the human cognition, while the models of communicative competence or language ability try to gain a comprehensive view of human behavior in the social context. Most of the latter work has been done in the context of language testing; even there, models that would deal with the affective or volitional aspects of human behavior are relatively recent (see eg. McNamara 1996).

To get an overview of the factors and processes involved in a task type such as sentence-level RT grammaticality judgments operating at the microlevel of language performance, a new model is needed. This model has to comprise various aspects of human performance. One such model was proposed by Sajavaara and Lehtonen (1980, see also Sajavaara 1981b, 1987b). In Sajavaara and Lehtonen's model, a dynamic view of language is adopted. They state that

> In addition to the dynamic nature of language acquisition [...]language use also entails dynamic aspects of human interaction and communication which should be integrated into the description of the phenomena involved in second/foreign language proficiency. (Sajavaara and Lehtonen 1980: 25)

To describe the factors involved in human communication, Sajavaara and Lehtonen suggested a model of message processing shown in Figure 10. They then defined communicative competence as an ability to make use of the mechanism described in the model; the term language proficiency was used also. They further pointed out that language proficiency is not the same as linguistic competence (Sajavaara and Lehtonen 1980: 30). Instead, Sajavaara and Lehtonen saw competence both as an ability and a "superior skill lying beyond the level of proficiency" (Sajavaara and Lehtonen 1980: 33).

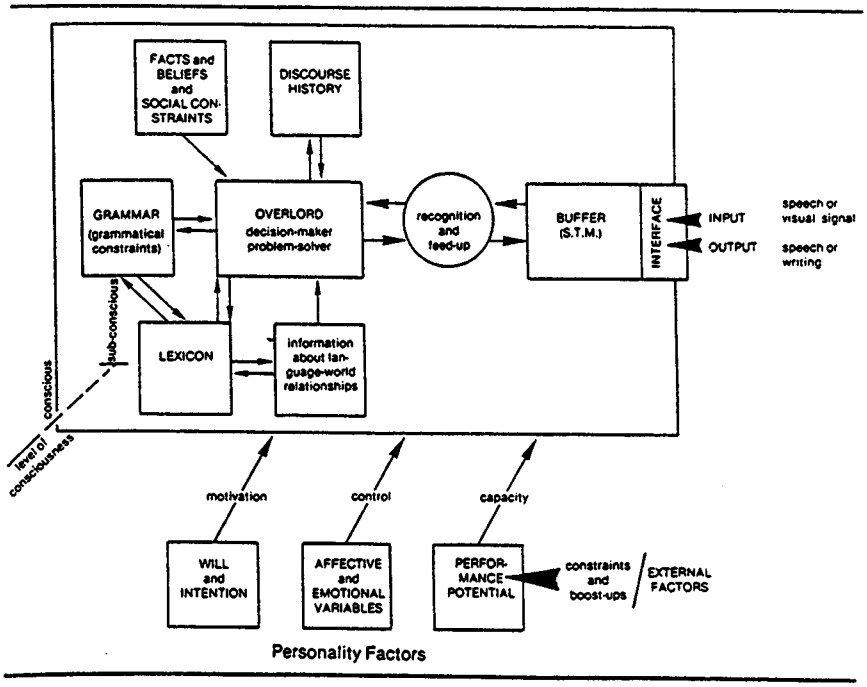


FIGURE 10 A simplified model of message processing (based on Sajavaara and Lehtonen 1980: 29).

Sajavaara and Lehtonen made the following points about the overall nature of the system:

> In perception, linguistic information is not processed 'hierarchically' by proceeding step by step through the levels of grammar from a concrete level of representation to phonology and syntax, and further to abstract 'meaning'. Instead, the components of the system which exploit the hearer's phonological, syntactic, and pragmatic knowledge work simultaneously in terms of a time-sharing system [...] there are 'cues' for all levels of grammar in the flow of speech waves and [...] the infomation involved is not present in the actual speech coded through phonology alone. (Sajavaara and Lehtonen 1980: 28).

The nature of knowledge and processing mechanisms involved are left unspecified in the model (although the nature of speech processing was discussed elsewhere (eg. Lehtonen and Hurme 1980). According to Sajavaara (1981: 99), the model is to be understood as a representation of the problematic areas in language processing and an attempt to "correlate linguistic elements to other elements" rather than a representation of the flow of data during language processing. For example, the arrows indicate relationships and channels for feeding in data and not the flow of data.

According to Sajavaara (1981:99), the model is based on a network of different types of knowledge needed for efficient communication, including not only grammar and lexis but also discourse history, facts and beliefs, and information about language-world relationships. The central feature in the model is the Overlord, a decision-maker/problem-solver, which 'governs' the knowledge network and has access to all the sources of information all the time. However, human communication does not involve only manipulation and transfer of knowledge. To include personality factors such as will, intention and emotion, Sajavaara and Lehtonen introduce three sets of intervening variables that affect the functioning of the system. These variables, represented by the three boxes outside the system proper, include will and intention, which contribute to motivation, affective and emotional variables, which influence control, and performance potential, which has its reflection in capacity (although

the precise relationship between these is not specified). These variables remain under the influence of external factors, which means that communicative performance is "never constant but varies under the influence of these factors" (Sajavaara 1981: 100).

The Overlord processes the information it has access to heterarchically, ie. in a parallel fashion, by comparing the incoming data and existing knowledge in order to make decisions as to the most probable interpretation of the message. It also makes corrections to earlier intepretations according to new information. The functioning of the Overlord is dependent on previous knowledge as it is needed to recognize incoming new information.

The third dimension in the model represents different levels of consciousness. According to Sajavaara, the nature of knowledge present in the system varies from

> fully conscious to totally subconscious: some data which is constantly used for purposes of interpretation never reaches the level of consciousness, some lies at the borderline, and some data is based on explicit formulations of rules and definitions (the borderline is never strict and the whole is to be seen as a continuum). (Sajavaara 1981: 101)

What Sajavaara and Lehtonen mean by consciousness is left to be inferred; as Schmidt (1990, 1994) points out, the term is notoriously indeterminate. However, it appears that the 'level of consciousness' included in the model refers both to the implicit-explicit nature of knowledge and to the individual's ability to consciously perceive it, ie., the individual cannot perceive consciously the processes involved; instead, "what he perceives are the final choices" (Sajavaara and Lehtonen 1980: 31).

The buffer, including short-term memory, and the interface divide the incoming flow of information into chunks which can be processed by the system. The interface performs at least part of this function on the basis of physical features or 'cues' of the incoming signal. Sajavaara and Lehtonen note that the interface and the buffer are necessary for both the reception and production of speech since both the reception and transmission of language and the control of articulatory and hearing organs and motor programs are carried out here (Sajavaara 1981: 101).

A perceived key element in the incoming signal activates the syntactic, morphotactic and pragmatic rules as well as various semantic networks. This information, together with the existing knowledge and the earlier analyses of the incoming speech-signal, are simultaneously available to the Overlord. They act as constraints which "open up or restrict potential choices in the construction of message content or in the production of speech" (Sajavaara 1981: 101).

Sajavaara and Lehtonen (1980: 30) noted that the interrelationship between lexicon and grammar is both intriguing and problematic. As suggested by Marslen-Wilson and Welsh (1978: 58), each lexical memory item (or 'word') could be defined as an intersection of a cluster of procedures operating over a range of cognitive dimensions. In language processing, a 'word' activates, for example, certain frequent and prefabricated phrases, word combinations, certain grammatical constraints, selectional restrictions, and certain semantic concepts and fields. Sajavaara and Lehtonen (1980: 31) further point out that a 'word' is used here to refer to a lexical item "which includes all the information concerning its use in sentences and its relationships to other words" and suggest that "it may be that entire language proficiency is in fact embedded in the memory in the form of such lexical items".

Sajavaara and Lehtonen (1980) do not discuss in detail the nature of knowledge and automatic and nonautomatic processes involved nor how the system would incorporate knowledge from more than one language. The model was primarily intended to complement

the static view of language and interference present in contrastive linguistics and as such could be used to gain insight in the cross-language processes; for example, they note that the model predicts interference in the detection of cue patterns, lexical identification and "reconstruction of phonological target structure" in cross-language communication (Sajavaara 1981: 102). In retrospect, and in light of the new findings on the workings of cognitive processes, one of the greatest drawbacks of the model is that it leaves unspecified the nature and interaction between knowledge systems and the role of attention and awareness.

Still, the model is one of the first attempts to describe comprehensively the interrelationship between cognitive and affective factors in human communication. It serves as the framework within which the RT grammaticality judgment task can be regarded. Consequently, such judgments are primarily seen as instances of problem-solving (see eg. Birdsong 1989a) and decision-making that vary according to automaticity and the use of more conscious problem-solving strategies.

3.6 A psycholinguistic model of second language metalinguistic performance

Grammaticality judgments are generally regarded as a type of metalinguistic performance (Birdsong 1989a, Ellis 1991). Linguistic performance can be described "in terms of the architecture of speakers' linguistic knowledge, and metalinguistic performance in terms of access to and retrieval of that knowledge" (Birdsong 1989a: 54). Also, metalinguistic performance is "contingent on factors of metalinguistic skill as well as knowledge of language" (Birdsong 1989a: 54). Furthermore, it is not likely that the knowledge invoked in a judgment test can be "uncontaminated by psycho-cognitive operations which, though perhaps totally unrelated to the recall of linguistic information, nevertheless influence the availability of that information" (Hedgcock 1993: 4). In other words, factors very much like ones described in Sajavaara and Lehtonen's model above, motivation, affective variables, etc., influence the outcome of metalinguistic tests.

The nature of knowledge invoked in second language metalinguistic performance and the processes by which that knowledge is accessed and retrieved is thus of considerable interest. Some of the factors related to the development of automaticity and fluent, skilled performance have been discussed above. Just as linguistic performance, metalinguistic performance is trainable and can be characterized in terms of skill learning (Birdsong 1989a: 53).

Levelt et al. (1977) present two problem-solving models for grammaticality judgment and paraphrasing processes. Figure 11 below shows a modified version of these models; paraphrasing has been left out.

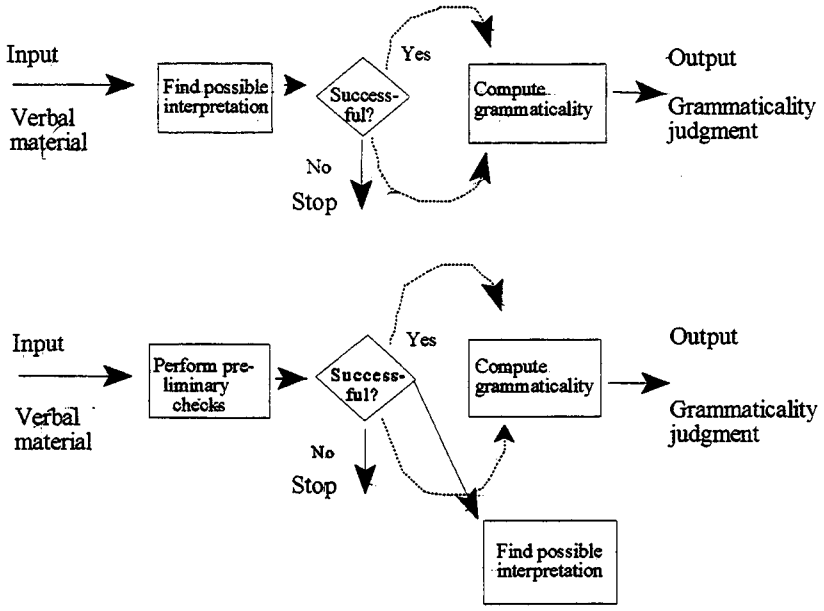


FIGURE 11 Two modified models of grammaticality judgment (based on Levelt et al. 1977: 98).

Levelt et al. (1977) present two alternative ways in which input can be processed: First, a possible interpretation for the item to be judged is found, ie. meaning is accessed first, and depending on the results, grammaticality of the linguistic item is computed. Second, preliminary checks are performed first, after which grammaticality is computed and a possible interpretation is searched for, simultaneously. Although our understanding of what goes into speech comprehension and production processes has deepened in the 20 years since these two models were proposed, they can be used as an overall guide into the processes involved in RT judgment tasks. These two routes may not be mutually exclusive either. On the basis of what we know of the development of automaticity and the parallel/serial nature of processing, it is possible that language users use both routes simultaneously to reach a grammaticality judgment (just as they use both routes in language processing). There are some findings that indicate that learners with a higher level of proficiency primarily use meaning-based strategies to judge the grammaticality of sentences (Ard and Gass 1987). (Native speakers have also been found to react differently to ambiguous sentences in RT acceptability judgment tasks, usually taking a longer time than language learners (Lehtonen and Sajavaara 1983).)

The two processes described here can be seen as a microprocess within a larger framework of second language performance. It is evident that there are several major sources of influence in the grammaticality judgment process. The first of them involves the type of knowledge the learner has access to.

Sajavaara and Lehtonen's model only specified knowledge in terms of its location along the consciousness-subconsciousness dimension. Birdsong describes language knowledge and its use in terms of Bialystok and Ryan's model of linguistic and metalinguistic functions. This model is one version of Bialystok's bidimensional model of second language proficiency which has undergone many revisions (eg. Bialystok 1990b). According to Bialystok and Ryan (1985a, 1985b), language performance can be described by its relationship to two dimensions of language proficiency, analysis of knowledge and cognitive control. Knowledge (ie., mental representations) can be unanalyzed or analyzed;

unanalyzed knowledge is typically implicit while analyzed knowledge is explicit and can be accessed in the course of problem-solving. Control refers to the ease of processing: how fast and easily knowledge could be accessed in different types of language use. It also involves three different functions: the selection of the items of knowledge, their co-ordinatation, and the extent to which selection and co-ordination can be carried out automatically. Attention plays an important role: an increase in control is correlated with an increase of the degree of attention that must be directed in order to succeed in a task (Bialystok and Ryan 1985a: 213-214). An increase in automaticity or fluency is reflected in a diminished need for cognitive control. According to Bialystok and Ryan (1985b: 235), the development of automaticity does not imply a change in the nature of knowledge from declarative to procedural as in Anderson's model. Furthermore, analyzed knowledge does not imply awareness or increased control; thus, analyzed knowledge can be accessed and used by automatic routines. In her later models, Bialystok limits the control to the "process of selective attention to alternative or competing representations in on-line processing" (Bialystok 1990b: 48).

Hulstijn (1990) contrasts Bialystok's model to the information-processing view on the acquisition of cognitive skills as described by Anderson (1983) and also to the view presented in Hulstijn and Hulstijn (1984). Hulstijn and Hulstijn (1984) suggested that the development of second language skills (to be exact, Hulstijn and Hulstijn limited their discussion to second language speaking skills) could be described by the dimensions of executive control and metacognitive knowledge (ie., knowledge of cognition). The dimension of executive control ranged from controlled to automatic processes, while the dimension of metacognitive knowledge involved the learner's metacognitive awareness, from highly implicit to highly explicit knowledge of language rules (Hulstijn and Hulstijn 1984: 25). Hulstijn and Hulstijn (1984: 25) hypothesized that "while controlling a structural feature, not all learners necessarily take recourse to an explicit linguistic rule in memory (the knowledge base) [...] some learners control a structural L2 feature with knowledge of an implicit, intuitive kind". Hulstijn (1990: 38) rejected his earlier view of two independent dimensions very much for the same reasons that he criticizes Bialystok's model for. Although the argument between Bialystok and Hulstijn (see Hulstijn 1990 and Bialystok 1990b) was mainly caused by terminological vagueness, it appears that the concept of automaticity remains indeterminate in Bialystok's model (cf. Schmidt 1992: 366-367). Also, Hulstijn questions whether it is necessary to have two separate dimensions for analysis of knowledge and control since it appears that both dimensions have come to incorporate both knowledge and skill. Finally, Hulstijn points out that the model can account for the development of metalinguistic skills (although, as he claims, not for the acquisition of primary linguistic skills such as listening and speaking).

According to Birdsong (1989a: 60), training can affect metalinguistic performance in at least two ways: First, training may reduce the need for higher levels of control. It is possible that because of increased automaticity, some individuals may perform better than others on tasks requiring higher level of control. Secondly, training can increase analyzed linguistic knowledge. For example, emphasis on the formal features of language, error correction etc., can lead to the development of higher levels of analyzed knowledge, and consequently to better performance on the tasks that require such knowledge. Bialystok and Ryan (1985a) list of a number of linguistic tasks as to their relation to the analysis/control dimensions of language proficiency. Acceptability judgments require only moderate analyzed knowledge; the more focus is on meaning, the less the task in question requires analyzed knowledge and cognitive control. As such, metalinguistic activities in general

require higher levels of control and/or analyzed knowledge than casual conversation and listening (Birdsong 1989a: 62).

Birdsong (1989a) presents a number of factors that may influence metalinguistic performance on a grammaticality judgment test, the most important of which are summarized below:

1. Repetition. Grammaticality judgments are not stable but tend to vary if repeated. Birdsong refers to an experiment (Birdsong 1984, in Birdsong 1989a: 64), in which 12 adult college-educated French L1 speakers were asked to judge the grammaticality of French sentences on a scale of 1 (not at all acceptable) to 5 (perfectly acceptable). Sentences were presented sequentially and the subjects were not permitted to go back to check items they had already judged. Two of the sentences appeared twice on the questionnaire. The subjects' responses for the second item were different from the first at least on one of the items. In a series of experiments, Nagata (1988, 1989b, 1989c) has shown the effect of repetition and context on grammaticality judgments.
2. Experimenter. The experimenter himself or herself may affect a subject's performance, for example, by giving examples of required responses, or otherwise affecting the subject's mindset by his or her own actions and words before or during the test.
3. Categorization. 'Grammaticality' and 'ungrammaticality' are usually considered two separate, although related, categories. A grammaticality judgment task can be regarded as a type of categorization task since in a grammaticality judgment task, a subject is typically asked to judge whether a linguistic item is a member of one or the other category. However, as is usual for example in the decisions concerning the membership of semantic categories, some items are felt to be "better", more prototypical members of the category. A related issue concerns the relationship between grammaticality and ungrammaticality: it is often viewed as a dichotomy; however, it can be viewed as a continuum as well with different degrees of grammaticality or ungrammaticality.
4. Validity. The results of grammaticality judgment tests cannot be generalized beyond the particular subject group or sample of sentences without at least some statistical evidence.
5. Order and mode of presentation. Greenbaum (1973, in Birdsong 1989a: 67) found that subjects showed a clear tendency to judge most severely those sentences presented early in the test. Learning may take place during very lengthy trials; fatigue may also play a role. When sentences are presented orally, intonation patterns may affect the judgments.
6. Response bias. Subjects may have a tendency to accept items containing relatively rare or complex structures (Quirk and Svartvik 1966, in Birdsong 1989a: 68).
7. Imagery and context. Subjects may often try to imagine a possible context in which the sentence they are asked to judge is acceptable. If this is the case, the nature of the items, ie. whether they are abstract or concrete, may affect the outcome of the test. Levelt et al. (1977) studied the effect of low vs. high imagery content in a RT grammaticality test of noun formations. Their findings suggested that items with high imagery content were accepted more often and faster than the more abstract items.

In his discussion of grammaticality judgments, Ellis (1991) lists a number of other variables that may affect the subjects' responses in a grammaticality judgment test: the use of distractor items, level of contextualization, what kind of response is elicited, how much time is allowed, and the familiarization with the judgment task. Subjects may also have developed response strategies; for example, they may attempt to strike a balance between judgments of grammaticality and ungrammaticality during a test. Hedgcock (1993) also discusses the linguistic and cognitive factors affecting second language metalinguistic performance. He divides the factors into two rough groups: linguistic variables, such as the

nature of knowledge the learner uses during a judgment test and the types of sentences presented, and extralinguistic variables, some of which have already been mentioned above.

Hedgcock (1993: 5) makes two important points: First, there is individual variation among learners in a grammaticality judgment test according to the access and use of explicit and implicit knowledge each learner has. Even learners with a comparable level of L2 proficiency may differ greatly. Second, as Ellis (1991: 164) points out, learners may base their judgments on an intermediate theory of language or on "what they think the target language norms are". Learners' interlanguage knowledge can be either implicit or explicit (or rather, more implicit or more explicit, if the distinction is viewed as a continuum). In other words, it is affected by the degree of automatization or internalization of L2 knowledge, or explicit knowledge made available during instruction. Quite often it is influenced by L1 and the perception of a relationship between the L1 and L2, as will be shown in Chapter 4. Also as Hedgcock (1993) points out, factors related to language processing may also affect the outcome of the judgments: for example, Schachter and Yip (1990) found that the learners' rejection or acceptance of highly complex or lengthy test sentences may have been due to processing difficulties rather than to the degree of ability to assess grammaticality based on L2 knowledge. (Of course, as argued by the present writer, if grammaticality judgments are regarded as tools for analyzing the *use* of L2 knowledge, such information in itself can be of great interest.)

There are also a number of task-related factors that influence the test performance, some of which have been mentioned above. The context of items may influence subjects' responses. The saliency and complexity of errors may affect the difficulty of test items (Galambos and Goldin-Meadow 1986, in Hedgcock 1993: 6).

Hedgcock (1993) lists a number of extralinguistic variables that may affect test performance. For example, learners may use test-performing strategies such as guessing. Some subjects may also have a tendency to reject an item whenever they are uncertain of its status (Bley-Vroman, Felix and Ioup 1988), and presumably vice versa. Subjects' motivation may also be important, not only as regards the taking of the test itself but also the way they view the importance of grammaticality and ungrammaticality, which may be the result of their learning experiences and objectives (Birdsong 1989b, in Hedgcock 1993).

To sum up the above discussion of the nature of variables that may affect a grammaticality judgment task, there appear to be three major sources of influence:

1. The nature of linguistic input, ie. the item to be judged. Ambiguous, grammatical (or well-formed, non-deviant, or correct) and ungrammatical (ill-formed, deviant, incorrect, or erroneous) sentences may be judged with varying accuracy and speed (if the latter is measured). Complexity and saliency of errors embedded in sentences may play a role, in addition to the context of sentences. Some sentences may be more prototypical examples of grammatical/ungrammatical sentences than others, which can affect learner performance, as well.
2. The judge himself or herself. Language knowledge systems, their nature, and various cognitive, affective and volitional variables play a role. Learners may have conscious or unconscious response strategies or biases. Learners may also change their approach to the task over time.
3. The test procedure and situation. The way the test items are organized and presented (for example, mode and order of presentation) may affect task performance, as well as such general variables as fatigue, noise, etc. The instructions and/or the demeanor of the experimenter may influence the test taker.

In the experimental part of the thesis, the findings of three different experiments will be discussed against the backdrop provided by second language performance in a RT grammaticality judgment task. The test sentences used in the experiments, in contrast to experiments conducted within the UG paradigm, were primarily based on errors typically produced by Finnish L1 learners of English at various stages of language learning. This approach, used in the Jyväskylä Cross-Language Project, receives indirect support from Hedgcock (1993), who notes that

> Few studies, however, have employed judgement tasks whose test strings consist in replications or close approximations of errors known to occur in the oral or written production of the learners in question [...] In cases where learners' productive use of the L2 is variable and not consistently accurate, the researcher may perhaps more readily tap into evolving interlanguage knowledge by eliciting intuitions concerning target language structures which have not yet stabilized. (Hedgcock 1993: 14)

Like Sajavaara and Lehtonen (1985: 7), Hedgcock (1993) suggests that there is a connection between production of a linguistic structure and intuitions concerning its grammaticality. Hedgcock (1992, in Hedcgcock 1993: 14) found that certain error types attested in adult subjects' L2 production yielded consistently inaccurate judgments on their part. Sutter and Johnson's (1990) study on L1 grammaticality judgments indicated that a low rate of accuracy in judging certain linguistic structures corresponded to a low rate of accuracy in producing the self-same structures (see Chapter 7 for further discussion).

4 Interaction between language systems: language transfer

A language learner has more than one linguistic knowledge system available to her. As the integral part of the present study deals with the use of grammaticality judgments to explore how the learner puts such knowledge to use during an on-line linguistic task under the circumstances in which interaction between the existing systems in encouraged, the next section will deal with the phenomenon of cross-linguistic influence, or language transfer.

As Gass and Selinker (1993; a revised version of the original volume published in 1983) note, the importance of some notion of language transfer had been understood within the field of second language research very early on. However, different theories of second language learning often hold very different views on the significance and even the very existence of this phenomenon. Thus, in discussing language transfer, the theoretical framework adopted by the researcher is bound to reflect on his/her views about transfer.

The term language transfer has also proved to be difficult to define. One of the reasons for this is no doubt the way language transfer has been and is still used to refer to a variety of phenomena from psycholinguistic processes to conscious strategies used in second language learning. Another term, cross-linguistic influence, which is sometimes used in second language research, avoids by its very vagueness any misunderstanding that might arise from the application of language transfer in different contexts. However, as long as it is borne in mind that language transfer has all the various senses mentioned above, there is no inherent reason against the use of the term; thus, the present study uses the term language transfer in the same sense that many researchers use the term cross-linguistic influence.

4.1 The definition of language transfer

The term transfer has its origins in learning psychology, where it means the effect of what has been previously learnt on the later learning (James 1980: 11). In behaviorism, the dominant approach to psychology in the 1950s, transfer referred to the transferring of learned habits. It began to be used in the study of second language learning to mean the influence of a previously learned language (usually native language) on the second language. As Robert Lado wrote in 1957:

> individuals tend to transfer the forms and meanings, and the distribution of forms and meanings of their native language and culture to the foreign language and culture - both productively when attempting to speak the language and to act in the culture, and receptively when attempting to grasp and understand the language and the culture as practiced by natives. (Lado 1957: 2)

During the years following these first pioneering analyses, the number of studies in the field of cross-linguistic influence proliferated. These studies tended to emphasize differences between native and non-native language systems (cf. Klein 1986: 25); they also attempted to discover the role that the language learner's mother tongue had in the learning of other languages. This approach became to be called contrastive analysis (CA).

The traditional claim of contrastive analysis went as follows: similarity between L1 and L2 meant that transfer from L1 and L2 resulted in successful performance. Transfer was then positive as it facilitated the performance. Structural difference between L1 and L2, on the other hand, resulted in an erroneous L2 expression. Transfer was then negative, and was usually called interference.

Based on such structural comparisons, Juhasz (1970; in Stedje 1977: 143-144), for example, distinguished between three kinds of transfer or interference. First, there is transfer which occurs when both languages have similar items. (This implies the existence of such similarities between the two languages that identification of for example phonological, semantic or grammatical structures can take place.) In such cases, the learner's L1 (or another language) could be of help. If there exist different structures which cannot be carried over directly, it is probable that L1 items will interfere with second language structures. Various instances of the use of 'false friends' belong here (Stedje 1977: 144). Thirdly, there may be no contrast between the source and target language, that is, there are structures in one language (such as the English article system) that have no equivalents at all in the other language, or that only one of the languages differentiates between certain items. As an example Stedje (1977: 144) mentions polysemy or homonymy as in the case of German *leicht* which can be translated with two words in Finnish (as in English): *kevyt* 'light', and *helppo* 'easy'.

There were indications that even in cases with no observable similarities between languages, L1 may have played a role in the formation of interlanguage. Schachter (1983), for example, pointed out that transfer may manifest itself also as more indirect, higher-order influence it has on hypothesis formation. The influence is then of a much subtler kind; for example, it could manifest itself as a tendency toward different types of hypothesis formation than native speakers of other languages might favour in the process of second language learning. This view of transfer led Schachter (1983) to include within its sphere any previous knowledge of language the learner may have, including the new information the learner has gained about the target language. Gass (1983a:80-81), one the basis of similar observations, makes a distinction between language transfer and transfer in language. According to her, transfer in language refers to L1-based phenomena in second language without any modelling pattern in L1 (cf. Dagut and Laufer 1982). Gass therefore limits the term language transfer to refer to instances of L1 patterns being transferred to second language. This kind of distinction is understandable, and necessary, if a view of language and of transfer is adopted which emphasizes the structural aspects involved (Lehtonen and Sajavaara 1985: 2). Usually transfer is easiest to verify when there is a clear source, such as a linguistic pattern in the source language, on which the interlanguage pattern has been modelled.

Thus, it turned out that the contrastive analysis view of transfer and interference was too simplified. It was rightly questioned whether differences in linguistic structures could be so straightforwardly equated with psychological processes and such subjective experiences as learning difficulty (see eg. Knapp-Potthoff and Knapp 1982, Sharwood Smith 1982). Along with the emergence of cognitive approaches to second language acquisition, the concept of transfer so central in the behaviorist school came under attack. The proponents of the L1=L2 hypothesis in the study of second language learning, such as Dulay and Burt (1974; see also Dulay, Burt and Krashen 1982), thought it necessary to show in morpheme studies that language transfer was not and could not be a significant factor in second language learning. Their view that there are other factors involved in second language learning which cannot be attributed to L1 influence is now actually widely

accepted (Gass and Selinker 1983: 6). On the other hand, the line of study proposed by these researchers and their followers proved to have serious shortcomings, as well. However, the damage to transfer research had already been done. The association with the behaviourist framework proved to be detrimental to the prestige of transfer and until the early 80s, transfer studies were regarded with certain reserve.

According to Selinker (1972), language transfer is one of the five central processes responsible for the development of interlanguage. In his 1972 article, Selinker does not give an explicit definition for language transfer; rather, he defines it through its relation to fossilization:

> If it can be experimentally demonstrated that fossilizable items, rules, and subsystems which occur in IL performance are a result of the NL [native language], the we are dealing with the process of language transfer;[...] (Selinker 1972: 216)

The difference between language transfer and a learning strategy remained unclear, as well. Some reseachers have later viewed transfer as a type of learning strategy (eg. Faerch and Kasper 1986, 1987). Language transfer can also be viewed as a phenomenon operating in on-line performance, ie. language processing (Lehtonen and Sajavaara 1985).

Language transfer is most often seen as a psycholinguistic process (see eg. Sharwood Smith 1982: 27; Lehtonen and Sajavaara 1985: 1; Odlin 1989: 161). This is in keeping with a general shift from product-based to process-centered approach in second language research. This has led to a redefinition of the term interference. As Lehtonen and Sajavaara (1985: 1) point out, the process of transfer cannot itself be called 'positive' or 'negative'; only the outcome, the product of such process can thus be named. Consequently, the term interference, as negative transfer is frequently called, should be restricted to the occurrence of errors in the learner's L2 (Kellerman 1977: 131). (Such distinctions are not always made: Dagut and Laufer (1982: 37), for example, refer to "the process of L1 interference".)

The complex and almost bewildering nature of language transfer is highlighted by other factors with which transfer seems to interact. For example, Odlin (1989: 28) makes a distinction between language transfer interacting with structural factors, by which he understands the comparisons made between two or more language systems, and nonstructural factors, such as social context, age, proficiency and personality of individual language learners. Furthermore, transfer's interaction with other processes of language production and reception, and of language learning (such as language universals or strategies of overgeneralization) must be taken into account (see R. Andersen 1983). Kilborn (1989: 1), working within the Competition Model, defines transfer as "the tendency for native language structures and the processing adapted to those structures to impinge on processing in a second language".

On the other hand, views have been presented according to which it might be better not to regard transfer as a process in itself. According to Ellis (1994: 28), L1 transfer usually refers to the incorporation of L1 features into the knowledge systems of the L2 which the learner is trying to build. He also points out the need to distinguish transfer as a learning process from the use of L1 as a communication process. As examples of L1 transfer as a communication process, he mentions translation, borrowing, code-mixing and code-switching. Ellis, however, notes the difficulty in distinguishing empirically between transfer as a learning process and transfer as a communication process. After all, a learning process can usually be studied only through products in situations where language is being

used. As we shall see, many linguists prefer to see transfer as a constraint on such linguistic processes.

Ringbom (1985a: 9) defines transfer as "that particular kind of reliance on prior linguistic knowledge which originates in the learner's L1". This type of definition, despite its vagueness, has the advantage of avoiding the difficulty in defining in detail what kind of process transfer is. It also avoids the difficulty inherent in the following definition proposed by Odlin:

> Transfer is the influence resulting from similarities and differences between the target language and any other language that has been previously (and perhaps imperfectly) acquired. (Odlin 1989: 27)

As Odlin (1989: 27-28) himself points out, the meaning of the terms 'influence' and 'acquired' as employed in this definition are not easy to determine. Furthermore, there are no adequate definitions for many other terms, such as strategy, process, and simplification (Odlin 1989:28). Thus Odlin (1989: 28) rightly concludes that any definition of transfer can be a working definition only, to be used until an adequate theory of second language learning can be formulated. Indeed, there is no positive need at this stage of research, when so little is known about the processes underlying language acquisition itself, to give any binding definition.

Today, the term cross-language (or cross-linguistic) influence is also used instead of transfer. As the behaviourist approach to language learning process fell into disrepute, this term has become to be preferred as one way of dissociating transfer research from the stigma that was attached to behaviourism (see eg. James 1980, Sharwood Smith 1982, 1986, Gass and Selinker 1993.) On the other hand, the term language transfer has at the same time become more and more distanced from its behaviouristic origins to the extent that many researchers feel quite safe using this term without being labelled as behaviourists (cf. Odlin 1989: 26). Similarly, to refer to the influence of learners' L2 on their L3, a term L3-interference has sometimes been used (eg. Stedje 1977). Nowadays, it is more popular to speak of cross-linguistic influence between non-native languages, or just non-native language influence. Ringbom (1985b: 39), for example, refers to this little-studied phenomenon shortly just as LN-influence (see also Ringbom 1986); in the present thesis, the term LNN-influence is used for non-native-language influence.

In this study, second language learning is seen mainly as integration of new linguistic knowledge with the already existing L1 knowledge system in such a way that the end result of the learning process is automatic, fluent performance in the target language. The role that language universals play in this process is left open. In such a model, declarative and procedural knowledge as well as automatization and use of automatic and controlled processes have an important role. The degree to which the speakers are conscious of their knowledge, ie. their level of consciousness (or 'awareness'), also interacts with the various levels of knowledge and processes.

Faerch and Kasper (1986,1987; see also Faerch, Haastrup and Phillipson 1984) have discussed various aspects of language transfer from the viewpoint of cognitive and psycholinguistic approaches to second language learning. According to Faerch and Kasper (1987: 115), transfer is related to declarative knowledge. This declarative knowledge can be either primary (in interlanguage communicative situations the primary declarative knowledge is learners' interlanguage knowledge) or secondary (learners' L1 or any other language knowledge they might possess). It may also represent various linguistic levels. Furthermore, it can be analyzed or unanalyzed (explicit or implicit) to varying degrees. Of

the types of knowledge, Faerch and Kasper (1987) regarded analyzed knowledge as most easily leading to transfer. Referring to transfer in production, they see it as

> that part of a learner's procedural knowledge which *selects* secondary declarative knowledge for the implementation of a subplan in the production process and *combines* this 'secondary' subplan with 'primary' subplans, so that the outcome is a primary plan, ie. an IL plan containing an L1 subplan (eg. phonological features, word order, semantic or pragmatic meaning etc.). (Faerch and Kasper 1987: 115)

As a reception procedure, transfer involves the interpretation of incoming L2 utterances by the language learner on the basis of his or her L1 knowledge. Thus, it represents a case of inferencing. (Faerch and Kasper 1987: 113).

4.2 Language transfer and language knowledge

Some linguists make a distinction between two kinds of transfer. According Sharwood Smith (1982), one linguistic system may influence another within the mind of a particular language learner in basically two different ways: First, the language learner's underlying knowledge, or his 'competence' (in the sense the term is used by Chomsky in 1965), may be affected. This refers to learners' current interlanguage, their transient competence. The network of relationships and categories that make up this competence may have been constructed in part by using parts of the L1 system. Secondly, there seems to be another kind of ability which is the set of mechanisms for activating underlying competence millisecond by millisecond, during language processing, ie. on real time.[1] This means that even if the learner 'knows' for example the right way of forming the word order of subordinate clauses in German, in the course of producing utterances under conditions of stress she may use her more efficient, automatized L1 routines; L1 processing habits will interfere with her performance (Sharwood Smith 1982: 32-33). Usually in this case the learner is able to correct the error if required.

These two types of transfer Sharwood Smith (1982: 33) calls competence transfer and performance transfer, respectively. Thus a distinction is made between the effects of the language learner's underlying beliefs (competence) and a lack of efficient processing mechanisms in second language. In 1982, Sharwood Smith pointed out that it is the learner's competence that should be the object of study (pp. 33-34). This view is understandable when one takes into account the framework of Chomskyan linguistic theory with its emphasis on competence as the only worthwhile object of study. By relegating transfer to the sphere of performance, those within the field of Chomskyan linguistics who were opposed to the idea of transfer sought to belittle it (cf. Sharwood Smith 1982; 1986: 245).

In his 1994 book on theoretical foundations of second language learning, Sharwood Smith (1994: 101) raises an important issue of the nature relationship between language transfer, as one of the interlanguage processes, and learner performance and competence

---

[1] Cf. the dictinction that Faerch and Kasper (1987: 112) make between transfer in learning and communication; transfer as a communication procedure takes place at the moment of speech production/reception.

(knowledge): to what extent is language transfer a process controlling performance rather than a process building knowledge? Here Sharwood Smith refers to Adjemian (1976), who was the first to raise the question of whether interlanguage, in the Chomskyan sense, belonged to the sphere of competence or performance, ie. whether it was primarily knowledge underlying performance behavior or the systematic aspects of that behavior.

Sharwood Smith (1994) argues that if language transfer (and overgeneralization) is a process that changes interlanguage knowledge, then interlanguage competence would contain 'copies' of L1 rules which were now applied to L2 material, although now part of a separate rule system. Moreover, "the learner's intuitions about the L2 (with regard to a given linguistic phenomenon) would be the same as the learner's intuitions about L1" (Sharwood Smith 1994: 101). For example, a French L1 learner of English L2 would regard a sentence *John stroked often a cat* as perfectly acceptable English because in their L1 a time adverbial would occur at the same place. Sharwood Smith uses knowledge to refer to 'mental grammar'; he points out, however, that it is unclear whether Selinker made a distinction between knowledge as 'mental grammar' and on-line processing mechanisms by which knowledge is used.

On the other hand, if language transfer has to do with on-line processing, ie. production and reception, during language performance, then learners would not take L1/L2 equivalence for granted. Their interlanguage system would not have 'copies' of L1 knowledge; instead, they would try to process L2 material through mechanisms originally set up to process L1 material. Consequently, transfer errors observed by the researcher or teacher would then be performance errors, not errors reflecting underlying knowledge.

Finally, Sharwood Smith (1994: 102) suggests that there is also a third possibility: that there is simply no valid competence/performance distinction to be made in Selinker's theory. Many linguists (eg. Hymes 1972, Halliday 1978) have presented this view, according to which interlanguage is simply 'systematic behavior', and that there is no difference between knowing a language and being able to perform in it.

There are other approaches to second language learning, as well, in which such a dichotomy need not necessarily apply. Thus, MacWhinney (1987: 135) sees the focus on the study of competence as "inappropriate" and "counterproductive". According to this view, the goal of second language instruction can be regarded as the development of a high level of actual language performance and not the establishment of abstract competence in the students. Consequently, the central concern of second language research should be the study of factors leading to changes in second language performance. The views presented by Bates and MacWhinney (1982, 1989) in the Competetion Model have produced a increasing number of studies on cross-linguistic influence in sentence processing (eg. Kilborn 1989, Kilborn and Cooreman 1987). This more than anything else is a proof of growing awareness of the importance of language transfer, whether it was regarded as a performance level phenomenon or not.

The relevance of the competence-performance dichotomy to the importance attached to language transfer seems also to be questioned in the cognitive theory described and advocated by McLaughlin (1987), in which transfer is seen to take place because the speaker has incorrectly activated an automatic routine based on L1 (cf. Sharwood Smith's performance transfer).

If one accepts the existence of competence and performance, and is mainly interested in competence as the object of study, there are, as Sharwood Smith (1986: 247) points out, some interesting methodological consequences. As the situations that put the learner under great stress may produce performance transfer, it is advisable not to investigate so much

transfer in production. Instead, one should examine transfer in situations where the learner's receptive language skills are used (Sharwood Smith 1986: 247). However, it is to be noted that even tests of this kind are susceptible to what Sharwood Smith calls performance errors.

Distinctions of a more operational kind are usually made as well. We can ask on what levels of language transfer occurs. Usually divisions will be made according to the level in which the errors induced by transfer occur. Consequently, we can talk of phonological, morphological, syntactic, lexical and discourse transfer. That it is not easy to draw a line between the phenomena on the various levels of language is not surprising if we think about the problems we have when we try to make such distinctions in linguistic analysis in general. In the study of transfer, a further difficulty is added by the fact that there are two languages involved, at least. As Stedje (1977: 150) notes, what is a phenomenon of the morphological level in one language is a syntactic feature in the other language, mentioning as an example the Finnish use of case-endings to express the same functions as prepositions in German and Swedish. However, as Ringbom (1982: 93) points out, the exact location of for example prepositions in the structural framework of language is likely to be irrelevant as long the focus of study is kept clear.

4.3 Errors and language transfer

L1- influence is shown most clearly when it leads to error (see eg. Sharwood Smith 1982: 27). Of course, transfer can also manifest itself as avoidance, overuse, and facilitation (Ellis 1994: 29). Transfer is thus usually operationalized through its manifestations in learner performance. To explain learner behaviour, the so-called strong version of the contrastive analysis hypothesis claimed that errors could be predicted on the basis of contrastive analysis (cf. Wardhaugh 1970). (Later on, the hypothesis was modified: now contrastive analysis only claimed to be able to explain errors (James 1980: 184-185).) However, in its attempt to predict learner behaviour, contrastive analysis soon ran into difficulties. This was mainly because its predictions, although probabilistic in nature, were extended to actual second language speech behaviour, even without careful descriptive and analytical studies of actual second language learners under clearly specified conditions (Gass and Selinker 1983: 3). Moreover, as noted above, the defect of traditional contrastive analysis also lay in its assumption that difficulty coincides with linguistic difference, and that the more errors there are in the performance, the more difficult the item in question is to learn (Kellerman 1977: 86). Difficulty does not necessarily have anything to do with error, though; correspondingly, a highly erroneous sentence may cause the learner no difficulty at all.

For such reasons, as Sajavaara (1983: 72) points out, a systematic analysis of learners' errors soon followed contrastive analysis: when contrastive analysis failed in the prediction of learners' errors, error analysis stepped in either as its substitute or even as its complement.

Corder (1967) was among the first to point out the importance of learners' errors: not only do they provide feedback to the teacher but they also serve as evidence of the language learning process itself and provide the learner with devices by which she can discover the rules of the target language. Errors are usually defined as the deviations from the norm of the learner's target language; their source can, naturally, be defined with much greater reliability if they are clearly based on a source language structure. As Ellis (1994: 51) points out, this definition raises a number of problems: First, this definition of error inherently

contains the notion of standard, the norm from which a particular linguistic form deviates; however, this raises the question which variety of the target language should serve as the norm. Secondly, a distinction can be made between errors and mistakes. According to Corder (1967), an error takes place when the deviation is a result of lack of knowledge, ie. competence, whereas a mistake occurs when learners fail to perform their competence. In other words, a mistake is the result of processing problems that prevent learners from accessing their knowledge of a target language rule, for example, through lack of automaticity, and, therefore, it is a performance phenomenon (Ellis 1994: 51).

Johnson (1996: 122) defines both of these concepts in terms of declarative and procedural knowledge. Errors arise from learners not having the appropriate declarative knowledge, or having false knowledge. Mistakes are a result of a lack of procedural knowledge, or processing ability. Both of them can be seen as one manifestation of a more general knowledge/processing distinction. Corder (1981) wanted research to focus on errors since they represented competence. (The latter involves a much greater controversy: should second language research focus on competence and is any other form of research useless? (See eg. Gregg 1990, Eckman 1994.) As Ellis (1994) points out, this distinction, outwardly so simple, hides a number of problems: First, there is a problem of identification: it can be very difficult to tell the two apart; secondly, it assumes that competence is homogeneous rather than variable. However, it is possible, according Ellis, that the learner's knowledge of the target form is only partial, ie. the learner has not learned all the contexts in which a form is possible. Thirdly, a disctinction should be made between overt and covert errors, ie. errors that are clearly deviant on the surface, and errors that appear in utterances that are superficially correct but do not mean what the learner intended them to mean. Fourthly, there is also the question whether research should explore only deviations in correctness, ie. rules of usage, or whether it should also include deviations in appropriateness, ie. rules of language use. The latter are generally referred to as pragmatic errors. Within the framework of this study, error identification should not present any problems: errors built in test sentences are in most cases overt deviations from the written, standard English core grammar. It is the identification of the learner performance in accepting or rejecting these sentences that needs to clarified.

After having been identified, errors can be classified in several ways: first, they can be classified according to their form or linguistic level on which they occur; thus, we speak of lexical errors, syntactic errors, errors in the use of articles, and so on. Second, errors can be classified according to their systematicity (see Corder 1974): presystematic errors occur when the learner is unaware of a rule in the target language and are random in nature; systematic errors occur when the learner has discovered a rule but it is the wrong one; postsystematic errors occur when the learner knows the correct target language rule but uses it inconsistently.

Thirdly, errors can be classified according to their likeliest source. For the most part, second language research has focused on psycholinguistic sources (cf. Taylor 1986). Psycholinguistic sources of errors are described in Figure 12.

Richards (1971), in an influential article, posited three possible sources for errors: First, there are interference errors which occur as a result of the use of elements from one language while speaking another. Secondly, there are interlingual errors which reflect the general characteristics of rule learning such as overgeneralization, incomplete application of rules, and failure to learn conditions under which rules apply. As a third source of error, Richards named language development, by which he meant the learner's attempts to build

ERRORS

competence ('errors')	**performance** ('mistakes')
- transfer	- processig problems
- intralingual (eg. overgeneralization, transitional competence)	-communication strategies
- unique (eg. induced)	

FIGURE 12 Psycholinguistic sources of errors (based on Ellis 1994: 58).

up hypotheses about the target language on the basis of limited experience. This third source remained rather vague and most researchers generally refer to two possible causes: interlingual errors caused by transfer from L1; and intralingual errors caused by the second language system (or second language interlanguage). It is possible to attribute errors to teaching methods (Selinker 1972; see also Sajavaara 1983: 72). In this case errors are usually called induced errors. Strategies of learning play an important role as the sources of the learner's errors; it was precisely this aspect of transfer that attracted the attention of the linguists interested in second language learning.

Transfer errors and intralingual errors can be further classified (see eg. Richards 1971, Lott 1983, see also Larsen-Freeman and Long 1991: 59). The problem is that it is often difficult to distinguish between transfer errors and intralingual errors, and even more so between various types of transfer and intralingual errors. As Kellerman (1983: 112-113) notes, such classification of errors may be immaterial since it is possible for one and the same form to have multiple psychological sources, which may further interact and reinforce each other. As a matter of fact, on the basis of linguistic form alone, no positive clues can be given about the origins of errors.

At the process level of language, however, transfer can, for example, be seen to be reflected in different degrees of automaticity in the language learner's processing of linguistic knowledge. For example, it can be assumed that both the language learner's L1 and target language knowledge become activated by linguistic stimuli during the process of language comprehension (Klein 1986: 13). Transfer occurs because the learner has incorrectly activated an automatic routine based on L1 (cf. McLaughlin 1987: 150). Such influence can be either inhibiting in nature, in which case the processing times are prolonged, or it can be facilitating (and negative, if it leads into error) (Lehtonen and Sajavaara 1985: 23).

Hulstijn and Hulstijn (1984) studied grammatical errors made by adult Dutch L2 learners during a story retelling task as a function of processing constraints and explicit knowledge. They made the assumption that while skilled speakers would have available ready-made subroutines for the monitoring and correction of errors, nonskilled speakers such as beginning or intermediate second language learners would not yet have fully routinized many of the structural rules of the second lanaguage. (The term monitoring is used by Hulstijn and Hulstijn (1984: 24) in the psycholinguistic sense to refer to review processes underlying such overt and covert behavior as editing and self-correction taking place during speech processing.) By varying attention (as in focus on form) and time pressure during the task, they found that learners without explicit rule knowledge gained from focus on form and absence of time pressure just as well as the learners with such explicit knowledge.

Applied to the grammatical acceptability judgment task, these results seem to suggest that the conditions inherent in the task should help learners during the task performance to focus on form; after all, learners' attention in such a task tends to be on form from the outset. The time pressure characteristic of a RT grammaticality judgment task, however, may have a dual role: On one hand, time pressure could influence learners' ability to detect deviancy from the L2 norm through inadequate perceptual processing and/or by forcing learners to rely on knowledge easiest to access; a typical example of such knowledge is L1-based knowledge. On the other hand, since the time pressure during a RT judgment task mimics the time pressure typical of an authentic language use situation, it may encourage learners to rely on their intuitive language knowledge rather than explicit rule knowledge. To sum up, a RT grammaticality judgment task may bring forth aspects of learners' linguistic performance that show the influence of automatized processes and the use of implicit knowledge than a paper-and-pencil grammaticality judgment task would do.

4.4 When does language transfer occur?

As Odlin (1989: 152) points out, language transfer occurs at all linguistic levels, such as morphology and syntax as well as phonetics, phonology and lexical semantics. However, it seems to operate within a set of constraints. One of the most intriguing questions in recent research is why language transfer does not always occur but, instead, seems to take place only on certain occasions and under specific circumstances. Furthermore, language transfer has had to answer to the challenge put forward by the discovery of developmental sequences in second language learning. The existence of such sequences, which have been discovered in the acquisition of negation, for example, indicates that at least certain grammatical subsystems of the target language seem to be learned in a more or less fixed order, irrespective of the learner's L1.

The following conditions, or constraints, on the occurrence of transfer have been put forward by various linguists who often represent quite different approaches to the study of second language learning, and to the role of transfer within it. Some researchers, for example, seek to limit transfer to a merely contributing role associated with such processes as overgeneralization and simplification (cf. R. Andersen 1983). An attempt has been made, however, to explain any inconsistencies or contradictory use of terms, if present.

First, it seems reasonable that two preconditions must be fulfilled before transfer can take place (Faerch, Haastrup and Phillipson 1984: 193-194): First, there must be a sufficient amount of formal and functional similarity between L1 and L2 with respect to the various linguistic levels and to the different components of communicative competence (cf. Wode 1981). Secondly, the learner must be willing to transfer (it is not quite clear in what sense the authors are using this word; it is unlikely that they refer to conscious willingness, though). In this process, she is affected by the following factors:

On the basis of his studies on how Dutch learners of English resort to transfer strategies, Kellerman (1978, 1983) has proposed that at least two constraints of differing nature seem to operate on language transfer. The first of them has to do with what

Kellerman (1983: 114) calls the perceived distance between L1 and L2.[2] This notion was first proposed by Kellerman in 1977. He noticed that Dutch students showed more reluctance to accept Dutch-based idioms in English than in German (Kellerman 1977: 125). This he looked upon to be a symptom of the operation of the students' idea of the relationship between the languages in question. In other words, the students were aware of the greater dissimilarity between Dutch and English and, consequently, were inclined to reject expressions that they felt were based on Dutch. That they showed more readiness to accept such idioms in German was, according to Kellerman, due to the perceived close relationship between German and Dutch. Results from other studies also seem to support such a view of language transfer. Thus, Finnish students transferred lexical elements into English, not from their L1, Finnish, but from Swedish, presumably because they considered Finnish to be too different from English (see eg. Sjöholm 1976).

The second type of constraint is called the "transferability constraint" by Kellerman (1983: 116-118). It seems that the perceived distance of two languages alone is not enough to predict cross-linguistic effects. In order to predict learner behaviour something more is needed: the potential transferability of the L1 structure has to be taken into account. By this he means the probability with which this structure will be transferred in contrast to other structures in L1. "Transferability" of a L1 structure is a theoretical notion based on native speakers' own perception of the structure of their language. As such, it is independent of the nature of the second language. The judgments of transferability seem to be remarkably stable across groups of learners despite differences in the level of proficiency, exposure, and age among the subjects.

Connected with the notion of "transferability" is the idea of psycholinguistic "markedness" of a linguistic feature. If the feature is perceived as infrequent, irregular, semantically or structurally opaque or in any other way exceptional, it is psycholinguistically "marked" and its transferability will be inversely proportional to its degree of markedness. An L1 rule must be perceived by the performer to be unmarked or potentially language-universal. In support of this Gass (1983a: 79) has shown that some aspects of language are more likely to be transferred than others, such as elements that are perceptually "salient" or semantically transparent (eg. the genitive form of the English relative pronoun). According to Kellerman (1983: 117-118), also linked to the ideas of transferability and psycholinguistic markedness are the notions of language-specific and language-neutral structures. Transferability is one factor that determines whether the L1 structure will be treated as language-specific (not transferable to a given second language) or language-neutral (ie. transferable to a given second language). These notions are, unlike the relative transferability of structures, dependent on the second language: certain structures will be ragarded as language-specific if compared with one second language; if the second language is different, the same structures may no longer be regarded as specific but language-neutral.

It seems that perceptions of specificity or neutrality are subject to changes over time, particularly as learners' metalinguistic sophistication, which thus also appears to play a part, increases (cf. Mägiste 1984: 303-304). The learner's metacommunicative awareness of L1 may indeed play an important part. As Kellerman (1977: 93-94) points out, the language learner may resort to transfer if he believes that there could be a relationship between the

---

[2] As Gass (1983a: 80) points out, there is no model of 'language distance' that would allow it to be empirically verified or quantified; therefore, most of the discussion tends to be intuitive in nature, though no less insightful.

native language and the target language at a given point. This implies that he is to a certain extent 'aware' of the similarities and dissimilarities between the native and target languages, even metalinguistically so. Although it is by no means clear what metalinguistic skills actually are, and what their role is in first language acquisition and second language learning, it seems that such element must be involved.[3] As a matter of fact, it is of particular interest within the framework of this study because of the method adopted in the experimental part: in making acceptability judgments; the learner is also likely to utilize her metalinguistic knowledge (whether it is applied consciously or not) as well.

Kellerman (1983: 122) also proposes a further criterion for transferability: in the absence of specific knowledge about L2, the learner resorts to what he sees as systematic, explicit, and logical in his existing interlanguage system, thereby treating L2 as a 'reasonable entity' (hence the name 'reasonable entity principle', ie. REP). Learners tend to transfer such L1 structures that conform to the reasonable assumption, and avoid transferring such structures that do not.

On the basis of such constraints, transfer can be predicted to occur when the perceived distance between L1 and L2 is small and when the structures involved are unmarked. Moreover, language learners are not likely to transfer such elements to their target language that would introduce irregularity and unsystematicity unless they know how to (Kellerman 1983: 129).

Kellerman's approach is to a certain extent related to that of Eckman (1977, 1985); Eckman's work is, however, within the framework of linguistic universals whereas Kellerman's notion is psycholinguistic in nature (cf. Faerch and Kasper 1987: 123). Eckman (1977, 1985) has proposed a specific Markedness Differential Hypothesis. According to this model, transfer occurs primarily where the L1 feature is unmarked and the L2 feature is marked. The hypothesis predicts that those areas of the target language that will be most difficult for second language learners will be those that are both different from L1 and relatively more marked. An interesting result from this approach is the notion of indirect effects between L1 and target languages that are typologically different. In such cases the resulting interlanguage may be a process rather than a product, ie. there may be an increased sensitivity toward certain solutions, a tendency to view language in a certain way, that influences the course of acquisition (McLaughlin 1987: 104; cf. Ellis 1986: 212; see also Schachter 1983).

On the basis of what has been said, it seems reasonably safe to make the following assumptions about Finns learning English (or other Germanic languages): It can be hypothesized that Finns as language learners have a highly pronounced view of their language as being dissimilar from English and other Indo-European languages precisely because it is typologically very different. Furthermore, Finnish language learners regard English as being closer to German and Swedish than to their L1. They may also tend to regard at least many of the formal features of their L1 as psycholinguistically "marked". In other words, their metalinguistic awareness of their L1 affects the transferability of such items. Needless to say, formal language teaching is a powerful medium in developing metalinguistic, and furthermore, formalized knowledge. It could be therefore predicted that in their judgments on English, German and Swedish, Finnish language learners would show relatively little direct transfer effect from L1 at least in the areas of lexis and morphology, probably also in syntax (with the possible exception of word order), but would probably show more errors based on the target language system itself; they can also be assumed to

---

[3] See eg. Birdsong (1989a) for a discussion of the subject.

be more willing to transfer from a typologically similar language, ie. English, to German (and Swedish.) On the other hand, it seems that the items of whose nature, "markedness", learners are not especially aware lend themselves more easily to the process of transfer. Semantics seems to be one such area, as well as the textual level.[4] Moreover, certain areas of language lend themselves more to conscious awareness than others. It is to be remembered that the perceived distance between languages is relative: the Dutch learners of English in Kellerman's study may have been unwilling to transfer items to English from their L1 because of the perceived dissimilarity of the languages; from the Finnish learners' point of view, Dutch may not be that far from English.

4.5 Non-linguistic factors affecting language transfer

As Ringbom (1982: 87) points out, there are a great number of variables involved in language transfer. Among the most important are the learner's age (adult language learners transfer more readily than children), mode of learning (a classroom context seems to favour transfer more than a natural environment), and stage of learning (transfer seems to occur more often at the early stages of learning) (Ringbom 1986: 159-160). Furthermore, various socio-psychological criteria such as the speech situation may also play a role in the learner's readiness to transfer (Faerch and Kasper 1987: 124; cf. Odlin 1989).

According to Taylor (1975: 88), beginning language learners rely more on transfer as a learning strategy than more advanced language learners. The results of LoCoco (1976: 61) also point in this direction. According to LoCoco (1976: 61), language learners' proficiency affected their readiness to resort to L1 in such a way that transfer occurred mostly at elementary stages of learning; the more proficient the learners were, the more they used the strategy of overgeneralisation from second language structures already known to them. For example, Stedje (1977: 146) noted that Swedish learners of German seemed to make lexical errors of the following type only at the elementary stage: **wir...pflocken* (*<plocka* (Swe.) + *pflücken* (Ger.)) ('we ...pick'); **alle haben Hatten auf* (*<hattar* (Swe.) (*Hüte* (Ger.)) + German plural ending *-en*) ('all have hats on') .)

On the other hand, language proficiency can also be seen as a necessary precondition to transfer (cf. Mägiste 1984). Ringbom (1982: 94; cf. also 1983: 207,210), for example, considers proficiency in the language from which transfer takes place to be important to the extent that such a language must be either learners' L1 or some other language in which they are fluent. Grammatical rule knowledge and semantic properties must be automatized before they can be transferred. Individual items, though, may be borrowed from a language the learner knows only superficially. According to Ringbom, transfer does not refer to such individual items but to different ways of combining items. A distinction can thus be made between transfer and borrowing.

The mode of learning may also influence the process of language transfer. Selinker (1972: 218) gives an example how the mode of learning in a classroom context can influence what learners transfer: As mentioned above, the methods used in language teaching play a part in the formation of the learner's interlanguage. Thus, the fact that a Serbo-Croatian

---

[4] Unfortunately, discussion of language transfer as it occurs at the discourse level of language is beyond the scope of this thesis.

learner of English uses the personal pronoun *he* in cases where he should use *she* even though his L1 makes a corresponding distinction between feminine and masculine, can be explained by the material and the method (especially if it is of the type that uses drills) applied in language teaching.

5 Individual differences in second language acquisition and processing

Most of the second language research that has been discussed in previous chapters has focused on discovering what the general processes are in second language acquisition and language processing; in other words, they aim at describing similarities among learners. This is in contrast with the study of individual differences, a field of study not as established in second language learning as it is in other disciplines such as psychology (Skehan 1989:1). Yet, as Skehan (1989) demonstrates, the investigation of individual differences in second language acquisition may contribute to a greater understanding of the factors involved. Carroll (1981:23), while talking about individual differences in language abilities, points out that if psychometric measurements are taken as an acceptable way of approaching language abilities, it can also be assumed that

> one can define, at least on a probabilistic basis, some common core of facts and rules about a 'standard' form of a language that are ordinarily learned or acquired together by educated speakers of the language; the verbal ability factor disclosed in psychometric studies measures, in effect, the extent to which the individual knows this common core of language facts and rules. (Carroll 1981: 23)

The investigation of individual differences has very much focused on the study of learners' language aptitude, motivation, attitude, learning strategies, personality and cognitive factors. Below, three of these cognitive factors, the cognitive style of field dependence/ independence, anxiety, and thinking styles will be dicussed in more detail as they pertain closest to the experimental part of the study.

5.1 Field dependence/independence

Field dependence/independence is one of the many cognitive styles mentioned in research literature. Cognitive style is a hypothetical construct used to explain the process of mediation between stimuli and responses (Goldstein and Blackman 1978:2). It has been defined in various ways (see eg. Harvey 1963, Witkin et al. 1971). Messick (1976: 6) described cognitive styles as "habitual modes of processing". Witkin et al. (1971) defined them as "the characteristic, self-consistent modes of functioning which individuals show in their perceptual and intellectual activities". A feature common to all theory and research on cognitive styles, however, has been an emphasis on the structure of thought rather than the content, ie. on how cognition is organized (Goldstein and Blackman 1978: 3). There are a great number of approaches to cognitive style such as tolerance of ambiguity, and reflection-impulsivity. All three cognitive styles mentioned here have been studied as to their relationship to second language acquisition. Naiman et al. (1978) in their study 'Good Language Learner' measured their subjects' tolerance of ambiguity and field dependence/independence, among other things. Hulstijn (1982) and Abraham (1983) have examined the relationship between the reflection-impulsivity dimension and the degree of monitoring.

Field dependence/independence is a construct associated with the work of Witkin et al. 1971, Witkin et al. 1977a, Witkin et al.1977b, Witkin, Goodenough and Oltman 1979). It involves a test of visual perception in which an individual's capacity to distinguish a given simple figure from a complex background is measured. This test is called the Embedded Figures Test (EFT) (Witkin et al. 1971). (Other measures include the Rod-and-Frame Test (RFT) and Group Embedded Figures Test (GEFT).) The better and faster a person is able to find the target figure embedded in a complex background pattern (ie., 'the field' in the name of the dimension), the more field-independent that person is. Field independent people are thought to be better able to separate figures from ground, the essential from the inessential and generally speaking be more analytic (Skehan 1989: 111). If a person does not perform well on the EFT, the person is classified as field-dependent. Such persons are considered more global, holistic in their perception of the world, and not so good at attending to it selectively.

The performance on the EFT is considered to be related to an underlying dimension of differentiation. Differention is a concept referring to an individual's capacity to distinguish gradations of a stimulus dimension (Goldstein and Blackman 1978: 175). On the basis of this larger dimension, the terms 'global' vs. 'articulated' have also been used to characterize the two opposite poles of this dimension (see eg. Widiger, Knudson and Rorer 1980, Bennink 1982). Although the contruct has usually been restricted to cognition, Witkin and his co-workers (see eg. Witkin et al. 1977a, Witkin et al. 1977b, Witkin, Goodenough and Oltman 1979) have extended it to personality differences. For example, field-independent people are thought to be more impersonal and nonsocial, whereas field-dependent people are more socially oriented (Witkin et al 1977b: 198).

Although the field-dependence/independence dimension originated in the field of perception, there has been extensive evidence that these styles are also in effect when the person is dealing with symbolic representation, as in thinking and problem solving (Witkin et al. 1977a: 8). Its relationship to language learning and learning in general has also been studied. Goodenough (1976), for example, reports on the results of concept-attainment tasks in which subjects were asked to distinguish between exemplars and non-exemplars of a class of stimuli. During the performance of such tasks, people often form hypotheses about the definition of the class concept, which can then be observed by using various techniques. The results revealed that when field-dependent people adopted a hypothesis-testing approach, their hypotheses appeared to be dominated by the salient cues, whereas field-independent people were better able to use the whole set of possible cues. In language learning, as Skehan (1989) reports, it has been proposed that field-dependent people, because of their greater social orientation, are more willing to communicate with others, which would lead to more verbal interaction, greater exposure to the linguistic input, and greater communicative competence, while field-independent people are thought to be better at analytic and cognitive restructuring. Consequently, in Skehan's view, both types seem to have their advantages in the second language learning situation. The cognitive style of field-independence has nevertheless been regarded by many (see eg. Carter 1988) as more effective, a fact that Abraham (1985), for example, attributes to a deductive teaching style that dominates second language classrooms (see also Bialystok and Frohlich 1978).

Naiman et al. (1978) found a significant correlation between French L2 learners' degree of field-independence and their performance on a French L2 oral production and a listening comprehension test. This was contrary to the original assumption which predicted that field-dependent students would do better on the more communication-oriented task. Similarly, Genesee and Hamayan (1980) discovered significant correlations between the EFT and performance on a general achievement test in French L2. In both of these studies, however, the

correlations, although significant, were not strong (between .31 and .35). J. Hansen and Stansfield (1981) studied the performance of approximately 300 Spanish L2 college students on tests of linguistic, communicative and integrative competence and related it to their scores on the field-dependence/independence test. The results showed that field independence was significantly related to the learners' performance on the linguistic and integrative competence tests. Again, the correlations were very low (between .24 and .43). d'Anglejean and Renaud (1985) reported that field independence was one of a learner characteristics significantly related to French L2 achievement by adult immigrants. Chapelle and Roberts (1986) found field independence to be a significant predictor of success on a number of linguistic tasks, including the TOEFL, a dictation test, a multiple-choice grammar test, and an oral communication test given to ESL students in an intensive English program. Jamieson (1992) looked at the relationship of reflection/impulsivity and field dependence/independence to the TOEFL scores by adult ESL learners in a university program. She found that field independence was positively related in a moderate degree to the scores on all the TOEFL subtests on listening, language structure, and reading. The correlations were not very high (.37 to .45), but field independence appeared to contribute more to language proficiency than the dimension of reflection/impulsivity. Other researchers such as Tucker et al. (1976), Bialystok and Frohlich (1978) and Ellis (1990) have not found any significant correlations between field-dependence/independence and language learning success (although Bialystok and Frohlich (1978) found that field independence was significantly related to foreign language aptitude as measured by the Modern Language Aptitude Test, MLAT). Thus, it appears that if there is a relationship between field-dependence/independence and success in second language learning, it is weak at best (Skehan 1989: 113).

The relationship of field dependence/independence to the second language acquisition process and processing has been studied, as well, although not to the same degree. Abraham (1983) investigated the relationship between various cognitive styles, including field-dependence/independence, and the use of the strategy of monitoring (as defined by Krashen) by adult Spanish L1 students in a university ESL program in different types of linguistic task. She hypothesized that among the written tasks (a fill-in-the-blank test, proof-reading, and composition), the most monitoring would occur on the fill-in-the-blank test and the least on the composition. No significant differences in monitoring were found on the three written tasks, however. Of the various cognitive styles, field independence was significantly related to the amount of monitoring on all written tasks. Again, the correlations, ranging from .31 to .37, were not strong.

Abraham (1985) examined the relationship between field-dependence/independence and the deductive vs. inductive example-based teaching style in a university ESL program. Teaching was carried out by two computer-assisted instruction programs, each designed to fit the conditions. The results showed that there was a significant interaction between field independence and lesson: field-dependent subjects performed better with the inductive lesson, while field-independent subjects got better results with the deductive lesson. It is to be noted, though, that the treatment consisted of one lesson only, which affects the reliability of the results. Carter (1988) investigated the relationship between field-dependence/independence and learners' performance on language tests in two different Spanish L2 programs, one stressing formal linguistic achievement, the other functional language proficiency. The results showed that field independent students were more likely to be successful in both types of teaching programs. This is in conflict with Bialystok and Frohlich's (1978) and Abraham's (1985) findings.

It has been suggested that the relationship that field-dependence/independence seems to have with thinking, problem solving and, at least to some extent, learning might be due to the fact

that field-dependence/independence is actually a disguised measure of intelligence (Skehan 1989: 114). Skehan bases this argument on a series of research conducted by J. Hansen and Stansfield (1981) and L. Hansen (1984); see also Stansfield and J. Hansen (1983). These studies demonstrated that, firstly, field independence may be related to the performance on some linguistic task types more than on others (such as the cloze-test); and secondly, that when scholastic ability was partialled out, the relationship of field independence to the various types of language tests ceased to be significant in most cases. McKenna (1984) has been highly critical of of the field-dependence/independence tests, as well. In his review of research, he comes to the conclusion that they are related to a measure of intelligence. d'Anglejean and Renaud (1985) also found that there was an extensive overlap between the subjects' scores on a field dependence/independence test and their performance on a measure of verbal intelligence. In addition, subjects' performance in the MLAT, to which field independence has been linked in some studies, has also been shown to be related to intelligence.

Nagata (1989a) studied the effect of field-dependence/independence on repeated Japanese L1 grammaticality judgments. He assumed that field-dependent and field-independent people might adopt differing cognitive strategies in grammaticality judgments presented under specific circumstances. Eleven field-dependent and fourteen field-independent subjects participated in the study in which they were asked to rate the grammaticality of sentences three times. After the first two times, the subjects received negative reinforcement. He hypothesized that the subjects who were field-independent would become stricter in their judgments after repetition, since they were expected to pay more attention to the structural properties of the sentences, while field-dependent people would become less stringent as they would try to find plausible communicative situations for the sentences. The study also investigated whether the subjects' response to failure, ie. negative reinforcement, would vary according to their cognitive style. Nagata's findings showed that negative reinforcement appeared to have no significant effect on the subjects' judgments, although field-dependent people tended to be more lenient after negative reinforcement. Repetition had an effect, though, as field-independent people turned out to be stricter in their judgments after repetition. Field-dependent people were not affected by repetition. The results of this study show further support for treating grammaticality judgments as a task type no different from any other task: As Nagata (1989a: 746) points out, the findings go against the generally held assumption in generative linguistics that grammaticality judgments are uniform among native speakers. The results also show that cognitive style may have an effect on problem-solving processes.

Skehan's (1989: 115) rather grim conclusion that the intelligence component of the field-dependence/independence test may account for the results could be quite well-founded. All such results have led some researchers (eg. Griffiths and Sheen 1992, Chapelle and Roberts 1992) to conclude that what tests like the GEFT measure is aptitude rather than a cognitive style.

5.2 Language anxiety

Skehan (1989) points out that in psychological literature, a distinction is made between general and specific anxiety. For example, a person may have a general anxiety trait which affects his or her behavior in all domains, or he or she may only feel anxious in certain kinds of situations, such as in communicating in public or in a foreign language. Research has also revealed that students

tend to react to anxiety according to their ability: high-ability students seem to benefit from anxiety, while low-ability and especially average-ability students perform poorly, or even fail (Skehan 1989: 115).

MacIntyre and Gardner (1994: 284) define language anxiety as the "feeling of tension and apprehension specifically associated with second language contexts, including speaking, listening, and learning". Most research into language anxiety has been conducted by means of questionnaires. These tend to focus on specific settings such as language teaching classrooms (eg. Gardner 1985) and the public performance aspects of language learning known as communication apprehension (eg. Lehtonen, Sajavaara and Manninen 1985; for the definition of the construct, see McCroskey 1984). Research has shown that language anxiety is negatively related to performance in standardized proficiency tests and course grades (see eg. Clément, Gardner and Smythe 1980, Gardner and MacIntyre 1993). Lehtonen, Sajavaara and Manninen (1985) studied foreign language communicative apprehension by Finnish L1 university students. They found that reliance on language proficiency, anxiety caused by errors, social anxiety and shyness, and reward obtained in the English classroom were the strongest factors relating to reticence and anxiety in L2 communicative situations. Ely (1986) investigated language class anxiety in relation to classroom participation and found that more anxious learners tended to participate less in classroom activities.

The generally negative effects of anxiety on second language performance have been explained through its effect on attention: it is thought that anxiety consumes attention and cognitive resources that could be allocated to performance (MacIntyre and Gardner 1989). Thus, anxiety can be connected with both the second language acquisition process and, if language achievement is measured by means of a test, the testing situation itself. The problem is how to distinguish between the specific types of language anxiety. Horwitz, Horwitz and Cope (1986), for example, have pointed out the need to find out more about the specific relation between language anxiety and second language acquisition and communication processes. Tobias (1986) has developed a model for analyzing the effects of language anxiety on the language learning process. Tobias divided the learning process into three stages, Input, Processing, and Output (cf. Chaudron 1985, Gass 1988, Ellis 1994). According to Tobias (1986), anxiety may affect attention, concentration, and encoding at the Input stage, resulting in a missing input (cf. Lehtonen Sajavaara and Manninen 1985: 61). Anxiety influences cognitive processing at the Processing stage by increasing the processing time on tasks that are more difficult, more poorly organized and more heavily reliant on memory . At the Output stage, learners are required to demonstrate their ability to use the L2, which is usually measured by test scores or verbal production. By inserting various tasks such as number recall at the Input stage, MacIntyre and Gardner (1991, 1994) have been able to get a more process-oriented view on the actual effects of language anxiety on the second language acquisition process. On the other hand, Skehan (1989), following Bailey (1983), suggests that approaches aimed at gathering more qualitative data, such as diary studies, might give a better insight into the different aspects of anxiety, especially when so many factors (level of language proficiency, age, stage of learning) seem to be involved.

To sum up, there appears to be a relationship between measures of anxiety and learning. However, this relationship is not very strong as most studies suggest a negative correlation of about -.30 (Skehan 1989: 118). As was pointed out above, there are a number of additional factors involved. The more process-oriented approach used by MacIntyre and Gardner (1994) may give a more detailed picture, though, on the real-time effects of anxiety on second language performance. The ease of administration connected with questionnaires, however, is lost.

5.3 Verbal and visual thinking styles

Verbal and non-verbal thinking styles are closely connected with the work of Paivio. Paivio (1971) developed the Individual Differences Questionnaire (IDQ) (published by Paivio and Harsman 1983) to gather data to test his Dual Code hypothesis on the role of imagery in cognition and language processing. This hypothesis states that information can be represented and processed using either verbal or imaginal modes and that individuals differ in their use of these modes. The verbal and imaginal systems are functionally independent, yet at the same time interconnected (Paivio 1978: 380). In other words, verbal and non-verbal information is stored in long-term memory in two functionally independent but interconnected systems (Bourne et al. 1986: 172). Paivio's Dual Code hypothesis has been criticized, partly because of its neobehaviorist nature and partly because of his view of the nature of symbols. Others such as Pylyshyn (1973) and Anderson and Bower (1973) have suggested that both verbal and non-verbal material is stored in the same abstract format. This view has come to be known as the propositional theory.

The IDQ has been used to measure individual differences in verbal and visual cognitive styles. Spitzer (1986) studied the relationship of visual and verbal thinking styles to the effect of a cognitive intervention in young college students and people over 60 years of age. It was hypothesized that verbalizing in a problem-solving situation affects people in different ways according to the individual's preferred mode of thinking; both age and sex-related differences were expected to appear. In other words, problem solving by individuals who were more verbal would be generally facilitated by verbalizing, while such intervention during problem solving would be detrimental to the performance of individuals who were more visual (Spitzer 1986:142). In psychological research, verbalization seems to have both a positive and a detrimental effect on problem solving. The positive effect has, among other things, been attributed to the effect that labeling has on attention as it tends to focus it on the specific features of an item or situation (eg. Furth and Milgram 1973). Verbalization may also help to encode information into memory in a way that makes items more available for retrieval (eg. Furth and Milgram 1973, Loftus, Miller and Burns 1978). The detrimental effect of verbalization has usually been attributed to the additional load that it imposes on an individual's information processing capacity. While Spitzer found very few sex- or age-related differences in thinking style among the subjects, the younger group showed some effects of verbalization, in particular among the subjects placed at the two extreme ends of the visual-verbal dimension. In other words, verbalization appeared to help those subjects who were extremely verbal and hinder those who were extremely visual. It appears, therefore, that there could be some connection, however tenuous, between verbal and visual thinking styles and verbal processing, although this remains very much hypothetical.

Some researchers have been interested in the relationship of imagery and L1 grammaticality and acceptability judgments. Levelt et al. (1977: 89) noted how informants, if asked to describe how they perform the grammaticality judgment task, usually answered something like: "I try to imagine a situation in which the phrase or sentence could be used." This seemed to indicate that informants tried to find a cognitive, preferably visual context for the sentence, ie. somehow to use imagery. Levelt et al. (1977), while referring to the opposing views of the representational system (see above), take no position regarding that controversy.

They (1977: 90) point out that the instruction to 'use imagery' has been shown to facilitate a large variety of comprehension and memory tasks. Levelt et al. (1977), however, link the effect of the high or low degree of imagery to items used in test materials: concrete material , because it is 'high imagery', could be easier to process whereas abstract material, being 'low imagery', might result in errors and longer reaction times. This claim was supported by the results of an experiment in which Dutch L1 speakers were asked to judge the grammaticality of Dutch concrete and abstract compounds and then paraphrase them. Thus they seem to regard individuals' use of imagery as a function of the linguistic stimuli they are exposed to rather than a pervasive mode of thinking.

6 Grammaticality judgments

6.1 Metalinguistic judgments

According to Birdsong (1989: 2), metalinguistic performance is most often understood as subjects' responses to experimental tasks such as paraphrasing an utterance or judging the synonymity of two sentences. A prototypical example of metalinguistic performance is the rendering of judgments of sentence grammaticality or acceptability. The difficulty that a great deal of UG-based second language research has had with accepting this view is of course the perceived difference between 'linguistic' and 'metalinguistic'. Second language acquisition research is, after all, supposed to focus on 'linguistic' rather than 'metalinguistic' since 'metalinguistic' presupposes a level of awareness, a level of potentially explicit consciousness that does not fit well with the acquisition-learning distinction prevalent in many theories of second language acquisition. Yet, it must be borne in mind that metalinguistic knowledge need not be explicitly conscious, either. Such views of language proficiency or second language acquisition that accommodate the transition of explicit knowledge into implicit knowledge, or make use of the distinction between declarative and procedural knowledge, or controlled and automatic processing, or the operation of cognitive processes on L2 input, or approaches that see second language acquisition as a form of skill-learning have it easier to accept the use of such judgments as a window to linguistic knowledge, or to see it as a tool to gauge the potential for performance which is available for the learner.

Thus, grammaticality judgments can be considered to be useful for the exploration of the grammatical system. As Bley-Vroman and Masterson (1989: 209) point out, judgment-giving seems to be an example of skilled performance but the variation discovered in the application of this skill (see eg. Coppieters 1987, Nagata 1987, 1988, 1989a, 1989b, 1989c) does not mean that the grammatical system itself does not exist. Birdsong (1989a), for example, has proposed that grammaticality judgments should be regarded as one type of decision-making or judgment-making behavior among many others behaviors, and as such they may be better described in terms of principles given by more general models of cognition.

6.2 Grammaticality judgments in L1 research

Chomsky (1965) was the first to propose that the object of a linguist's study should be language competence, as opposed to language performance, of a native speaker. Chomsky (1965: 4) defined language competence as "an ideal speaker/hearer's knowledge of his language". Chomsky claimed that part of a native speaker's language competence was his ability to make judgments about the grammaticality of sentences, and that grammaticality judgments reflect the speaker's competence. This has served as a theoretical basis for the use of grammaticality judgments in L1 and L2 research (see eg. Mazurkewich and White 1984). For example, speakers were asked to judge the acceptability of such sentences as **John was easy for Mary to expect to come to the party.*

Not surprisingly, the notion was born that if a native speaker was asked to judge the grammaticality of sentences in his or her native language, the reaction time latency thus achieved would somehow reveal the actual stages in sentence parsing that take place in the speaker's mind. Neither was it a surprise that researchers soon discovered that reaction time measurement used in this way was too crude a tool for such purposes.

Grammaticality judgments have been used in L1 research, both as such and in combination with reaction time measurement, to study normal and abnormal child language acquisition and adult aphasia, in particular (eg. Lukatela, Crain and Shankweiler 1988, Wulfeck 1988, Wulfeck and Bates 1991, Wulfeck, Bates and Capasso 1991, and Blackwell, Bates and Fisher 1996).

In child language acquisition, researchers try to elicit this type of data from children to gather data on the stages of cognitive development and paths of approximation to adult language (Birdsong 1989a: 2). Ryan and Ledger (1979) studied the developing awareness of L1 syntactic structure among 5-8-year-old children through sentence judgment, repetition, and correction tasks. Judgments of grammaticality and correction improved significantly across the first years in school. The measures were significantly interrelated; they also correlated with reading scores. Training did not, however, appear to have any effect on grammaticality judgments. Ryan and Ledger concluded by suggesting that training and correlational studies regarding the link between syntactic sensitivity and beginning reading might best focus on the ability to attend to form in order to extract meaning rather than the ability to attend to form for form's sake.

Sutter and Johnson (1990) studied the ability of children, ranging from 6 to 8 years of age, to monitor grammaticality in the past progressive, perfect progressive, and perfect verb forms. It turned out that the children achieved a significantly higher rate of accurate judgments monitoring grammatical forms than ungrammatical forms. The results could be explained by the context surrounding ungrammatical verb forms and children's age, which appeared to significantly affect monitoring.

Nagata (1987, 1988, 1989b, 1989c,), in a series of experiments, studied the effect of repetition on grammaticality judgments by adult Japanese L1 speakers. Nagata's (1988) findings indicate, among other things, that judgments of ungrammatical sentences were more lenient when the sentences were embedded in context, rather than presented in isolation. Furthermore, it also appeared that linguistic intuitions as revealed in grammaticality judgments were not absolute but relative in that they were easily influenced by repetition and other variables, such as embedded context. Nagata (1989b), however, found no shift in judgments after repetition, calling the validity of his earlier experiment into question. Nagata (1989c) conducted three experiments with 60 Japanese university students to investigate the differential effect of repeated presentation of sentences on judgments of grammaticality under the two mental states of objective and subjective self-awareness. His findings seemed to indicate that repeated exposure to sentences exerted differential effects on grammaticality judgments dependent on whether the subjects were objectively or subjectively self-aware. All in all, Nagata's findings show that grammaticality judgments are not immutable but subject to change, even among native speakers. Moreover, as Nagata (1989a) demonstrated, grammaticality judgments may vary over time according to cognitive factors such as cognitive style.

6.3 Grammaticality judgments in L2 research

6.3.1 Methodological considerations

As in the case of L1 competence, it has been commonly assumed that metalinguistic performance reflects interlinguistic competence (Birdsong 1989a: 4). Data from L2 learners have been used to determine the effect of language transfer on learners' interlanguage grammar (Lehtonen and Sajavaara 1985; Birdsong 1989a: 2). Grammaticality judgments have also been widely used to study which sentences are possible in learner grammar (Gass 1994; see also Cowan and Hatasa 1994). The notion of learner language, of course, also entails the rethinking of the notion of language competence. This notion had come under heavy criticism from various sources during the 1970s. Furthermore, it was questionable in the beginning whether such a notion would even apply in second-language learning. It was even asked whether L2 learners can have intuitions about their L2; the answer was clearly yes (Schacter, Tyson and Diffley 1976). Regardless of whether L2 learners could be said to have 'language competence', they clearly had opinions about the grammaticality and/or acceptability of L2 expressions. Such opinions were based on their L2 knowledge.

It is important to note that there is a link betwen the method used to collect data and the type of data obtained by using this method. Hyltenstam (1983) studied learners by using different techniques to gather data. These techniques included elicited production, manipulation of given linguistic material such as sentence combining and sentence completion, intuition and grammaticality judgment tests, and introspection, among others. His findings revealed that there was a great deal of individual variation in the performance of individual learners across different types of tasks. Hyltenstam concludes that different language data elicitation techniques may be appropriate for different linguistic phenomena and for different learners at different phases of acquisition. The choice of appropriate elicitation techniques depends on factors such as proficiency level, linguistic phenomena in question, age, and educational background. Hulstijn (1989) also points out that there is a link between the type of task used to gather data and the type of task used to teach that particular linguistic point.

Grammaticality judgments have been used in L2 research since the late 1970s. Chaudron (1983) has conducted an extensive review of research in metalinguistic judgments up to the early 1980s that includes other types of judgments as well. Ellis (1991) in his review of research done in learners' grammaticality judgments notes how popular research in learners' metalinguistic judgments has become.

Following Chaudron (1983), Ellis (1991: 162) divides metalinguistic judgments into two broad categories: judgments on grammaticality and sentiments about language and its use. In a grammaticality judgment task, the learner typically has to decide whether a sentence is well-formed or deviant. After comparing the learner's response to that of a native speaker, it is possible to evaluate the correctness or incorrectness of such a judgment. There are other types of metalinguistic judgments: the learner can also be asked to report on his attitudes or opinions about language or language use; he or she can also be asked to match sentences or identify the agent in a sentence, a type of metalinguistic judgment that has been used in second language research involving the Competition Model.

Ellis (1991) makes an important distinction between two types of grammaticality judgment tasks: First, there are tasks that require the learner to discriminate between well-formed and deviant sentences. Secondly, there are tasks that require the learner to either locate errors in

deviant sentences and correct them, or provide grammatical descriptions of errors. These two types of tasks differ in the nature of response that they require: it is possible to respond to the first type of task on the basis of intuition alone. This relates to whether the learner makes use of implicit or explicit knowledge of the second language (cf. Bialystok 1979; Gass 1983b). The execution of the first type of task also does not require verbalizable knowledge and only involves recognition not production processes. Sharwood Smith (1994: 80) draws a further distinction between tests of linguistic intuition and explicit metalinguistic judgments such as explaining an error and why it occurred. Sharwood Smith also suggests that judging the acceptability of a linguistic item is metalinguistic only in the minimal sense that subjects are asked to treat a word or utterance as an object; in other words, such judgments should be regarded as implicitly metalinguistic.

According to Ellis (1991: 162), both practical and theoretical considerations have been used to justify grammaticality judgment tasks. Such tasks usually provide a relatively easy and effortless way of collecting L2 data. The theoretical rationale has been that such tasks enable the researcher to investigate the learner's linguistic competence. They are seen to provide a relatively direct window into competence. Ellis quotes Arthur (1980: 180): "learner judgments of acceptability are a reflection of that learner's competence in the target language". This claim has, of course, its roots in the procedures used by descriptive and theoretical linguists. This has increasingly been called into question. As Ellis (1991: 164) points out, it can be equally well argued that a grammaticality judgment task is a performance task similar to any other kind of task. Gass (1994: 307) points out, however, that researchers within the UG paradigm do not claim that grammaticality judgments provide a direct access to abstract respresentations of linguistic knowledge. On the other hand, grammaticality judgments can give us information about what are possible and impossible sentences in the learner-languages in a way a production task may not do.

During a grammaticality judgment task, learners' responses may be influenced by outside factors such as fatigue, impatience, or a particular type of test-performing strategy (for example, not judging too many consecutive sentences as ungrammatical), or they may or may not try to access explicit knowledge.

A great problem in a typical grammaticality judgment task is that there is no way to get absolute certainty on what subjects' judgments are based on. They may or may not focus on the structure that is the target of the investigation, or they may respond to factors such as semantic anomaly or style of the sentence. The effect of context, or lack of it, in the processing of sentences has also to be taken into account. The source of knowledge that they base their judgments on is not clear either, as they may draw on their existing interlanguage or on what they think the target language norms are. They may even be affected by parsing problems: the sentence may pose them such processing difficulties that they judge it as incorrect even if it were a grammatical sentence. Examples of this are provided by Schachter (1989), who noted that both native speakers and learners consistently rejected sentences of the type *Which book did you say John believes offended many people?*.

There can be great between-learner and within-learner variation in grammaticality judgments. This has caused the reliability of second-language grammaticality judgments to be called into question (see eg. Coppieters 1987, Nagata 1987, 1988, 1989a, 1989b, 1989c). Subjecs who are similar in many other respects still display as much variability in their metalinguistic judgments as in their L2 productions. A learner may judge a structure as correct on one occasion but a similar structure as incorrect on other occasion. Tarone, Gass & Cohen (1994) discuss this issue comprehensively. Hedgcock (1993) has also examined the effect of well-formed and ill-formed strings in L2 metalinguistic tasks.

Furthermore, while discussing the notions of grammaticality, acceptability and appropriateness in terms of communicative success in L1, Enkvist (1992) points out that since grammaticality can be seen as a relation between an utterance and a grammatical description of its language, acceptability must be seen as the relation between an expression and a set of judgments produced by a group of informants. He poses the question whether acceptability can be measured without any attention to the type of person who makes the judgments or the situation in which the judgments are made.

Finally, Ellis (1991: 165) raises a question that is very important for the purposes of the present thesis. If metalinguistic judgments do not provide direct evidence of competence but elicit performance data, what kind of behaviour do they entail? Do they mirror those obtained from the analysis of production data? Arthur (1980: 182), for example, claims that "sequences the learner judges as acceptable will belong to roughly the same set as those he or she uses when speaking the target language". Chaudron (1983) lends support to Arthur's claim. Referring to his 1992 study, Hedgcock (1993: 14) also suggests that there may be a connection between adult L2 learners' intuitions concerning the grammaticality of linguistic forms and their production in speech. Sutter and Johnson's (1990) study on children's L1 grammaticality judgments indicated that a low rate of accuracy in judging certain linguistic structures corresponded to a low rate of accuracy in producing the same structures. Gass (1983b), however, refers to research showing inconsistency between learners' actual production and what they say that they do in a metalinguistic task (see also Tarone and Parrish 1985). Ellis (1991: 166) goes so far as to state that there is ample evidence to suggest that learners perform quite differently in a grammaticality judgment task and in oral production (of course, it can be claimed, as Sharwood Smith (1988) does, that such differences are due to the increased processing demands of the production task). Ellis, in fact, refers to two recent studies that seem to indicate that at least for beginning learners, judging the grammaticality of sentences and producing L2 structures constitute different and unrelated types of behaviour.

Ellis (1991: 181) suggests that more study will be needed to gain a fuller understanding of the grammaticality judgment test as a task type: "Researchers need to acknowledge that metalinguistic judgments constitute performance data and be prepared to investigate what kind of performances are involved - much more than has been the case in the past."

6.3.2 Review of L2 research

Grammaticality judgments are quite extensively used in the study of second language acquisition in basically two ways. First, there are studies that use them to measure the development of a grammatical system in the acquisition of artificial grammars (eg. Reber 1976; 1989; Reber, Allen and Regan 1985, Dulany, Carlson and Dewey 1984, Perruchet and Pacteau 1990, DeKeyser 1994, 1996). They have also been used to study natural languages. In a great number of experiments, grammaticality judgments have been used to ascertain the level of L2 development or the state of L2 competence as to the occurrence of a particular grammatical feature.

Masny and d'Anglejean (1985) investigated the relationship between L2 grammaticality judgments and various cognitive and linguistic variables. These variables included L2 proficiency, L2 classroom achievements, L1 reading competence, language aptitude, nonverbal intelligence, field-dependence-independence. Analyses showed that L2 proficiency, L2 achievement in the classroom, and language aptitude were significant predictors of the subjects'

ability to make grammaticality judgments. L1 reading competence was significantly related to the subjects' ability to correct deviance.

Ard and Gass (1987) used a grammaticality judgement test to investigate the relationship of English L2 proficiency to syntactic acquisition. Four syntactic structures were chosen for examination. The results suggested that the less proficient subjects used syntactic strategies, while the more proficient learners used semantic-based strategies.

Johnson and Newport (1989) investigated the effect of critical language learning period on second language acquisition. They used a grammaticality judgment task involving errors in infinitive structures, 3rd person singular *-s* and article usage, among others to study the English L2 proficiency of Asians who had moved to the United States at varying age. The results supported their hypothesis as the early arrivals were significantly superior to later arrivals in English proficiency.

Bley-Vroman, Felix and Ioup (1988) investigated whether Universal Grammar (UG) is accessible to adult language learners by using a grammaticality judgment test given to advanced Korean L1 learners of English. Hirakawa (1989) used a multiple-choice grammaticality judgment task to study the acquisition of English reflexives among native speakers of Japanese from three different age groups ranging from 15 to 19 years of age. Hirakawa worked within the UG framework; her findings clearly revealed, though, the influence of learners' L1 on their judgments. Schachter and Yip (1990) studied the processing of English sentences with varying object and subject structures by native and nonnative speakers of English. It appeared that the native speakers showed a processing preference. A possible L1 influence was detected since a group of nonnative speakers whose L1 grammar had a certain structure also exhibited a tendency to process sentences in a certain way.

Inaba (1993) studied the acquisition of Japanese time adverbial clauses by native speakers of English. She used a grammaticality judgement test involving correct Japanese L2 sentences that were compatible with the English L1, ones that were incompatible with the English L1, and incorrect Japanese L2 sentences that could occur if learners transferred the English L1 grammar to the Japanese L2. She found that even in cases where the learners had positive evidence of the occurrence of a structure in Japanese that had no counterpart in their L1, the learners rejected it. Moreover, the learners accepted the L1-based structure in Japanese although there was no positive evidence to allow for it. Inaba concluded that the learners appeared to postulate L1 grammar as an interim theory about the L2, and that this tendency was remarkably strong especially at the initial stages of L2 acquisition. Similar results have been gained from transfer research (see eg. Taylor 1975, LoCoco 1976). Moreover, Alanen (1995) found a tendency by beginning English L1 learners of semi-artificial Finnish L2 to produce an incorrect L1-based structure in L2 despite positive evidence to the contrary, in particular among learners who had received explicit rule-based instruction in the L2 target structures.

Hedgcock and Lefkowitz (1993) investigated the perfomance of French L1 learners of English as a foreign language on a grammaticality judgment and preference tasks. The subjects were asked to utilize second-language knowledge brought to mind under an aural priming activity or a written task. Towell, Hawkins and Bazeragui (1993) conducted a longitudinal study to investigate variability in learner language by using a grammaticality test.

Winitz (1996) conducted a study on whether the methodologies of explicit and implicit language instruction could account for differences in the identification of grammatically well-formed sentences for college students of Spanish L2. For the explicit condition, a grammar-translation approach was used in which students were taught explicit statements of grammar rules. The students in the implicit classes were taught in the comprehension of Spanish L2 sentences

by using pictures and Total Physical Response activities. Interestingly, the results showed that the students who had received implicit instruction scored significantly higher in a grammaticality judgment test than those who had received explicit instruction. This supports Sharwood Smith's suggestion that acceptability judgments (the term that he prefers to use) are only minimally metalinguistic, and that successful performance on this type of task may not be helped by explicit metalinguistic knowledge. However, as Winitz (1996: 32) points out, the terms 'explicit' and 'implicit' are used in the study to refer primarily to language teaching approaches. It raises the question, therefore, whether the student in the 'implicit' class really learned implicitly (cf. Alanen 1995). It is more likely that this is a case of incidental learning through an inductive approach.

Ayoun (1996) investigated the acquisition of the oblique case by L2 learners of French. More specifically, she studied within the UG-framework, the applicability of the Subset Principle in the acquisition of the Oblique-Case Parameter, with the aim of determining whether the Oblique-Case Parameter really is a parameter of Universal Grammar. Since the subjects failed to reject a number of ungrammatical sentences in a grammaticality judgment task, she suggested that L2 learners may need direct or indirect negative evidence to constrain their grammar.

This gives only a brief overview of the many studies using grammaticality judgments as a data collection tool. In many studies, they appear to have served a useful purpose by revealing L1 influence, for example. Provided the methodological limitations and caveats involved in the use of grammaticality judgments are borne in mind, the results obtained by using them can be intepreted as any other language data.

7 Grammaticality judgments and reaction time measurement

In discussing the state of grammaticality judgment research, Ellis (1991: 181) points out that there is a need to obtain information about what learners actually do when they make judgments by, for example, timing their responses or by obtaining introspective information about learners' thought processes. There has been a great deal of research done in this area lately. Much of the research done in the 1980s and 1990s in Jyväskylä in connection with the Finnish cross-linguistic project is in a sense what Ellis recommends, that is, timed learners' responses in grammaticality judgment tasks.

7.1 Reaction time measurement

Reaction time measurement is a tool widely used in psychological and neuropsychological research. The basic hypothesis has been since the days of Donders in 1868 that reaction time could be used to estimate the speed of internal cognitive processes. Reaction time measurement has also been extensively used in the study of language processing. Usually such tasks have measured choice reaction time, ie. subjects have been asked to react in either of the two (or more) ways. For example, they have had to decide whether a sentence they heard or read was grammatically acceptable, or whether a pair of grammatical/ungrammatical sentences matched. Reaction time measurement can in fact be combined with any number of tasks; for example, it has also been used as a supplemental on-line measure of language performance in agent-decision tasks typical of the Competion Model, as in the work of Kilborn (1989).

Smith (1968, in Lachman et al. 1979) conceptualized such a choice reaction time (CRT) task as a sequence of mental events intervening between CRT stimulus presentation and response. The sequence had four stages:

1. The raw stimulus is preprocessed, making a clear representation for later processing;
2. The stimulus representation is compared with items in memory until it is categorized;
3. The categorization is used as a basis for response selection;
4. The subject then programs his response execution.

As Lachman et al. (1979: 150) point out, Smith's CRT framework illustrates how cognitive psychologists can use this strategy to understand people's information processing. Freedman and Foster's (1985) work on sentence matching is based on an assumption that the language system produces a high-level representation of grammatical sentences, which is then used as a basis for comparison. Matching ungrammatical sentences takes longer than matching grammatical ones, either because a high-level representation cannot be made and the matching must be done by using slower strategies, or because not one but two representations are constructed: the ungrammatical sentence as it is and a corrected version of this; it is the process of correction that prolongs the reaction time (Bley-Vroman and Masterson 1988:214-215).

However, it needs to be pointed out that there is no generally accepted theory or model of choice for reaction time yet (Jensen 1985: 54). Nor is there a single accepted way of defining reaction time. For example, reaction time has been defined as the minimum amount of time needed for the observer to produce a correct response. The problem with this definition is, as Jensen (1985: 67) points out, how we define what is the minimal reaction time of a given subject. Jensen further points out, however, that all that is important for a definition of reaction time is that it be made explicitly operational in terms of the details of the experimental paradigm that is being used to measure the time needed for reaction.

According to Lachman et al. (1979: 180), it appears that there is a set of pattern-recognition processes that occur very early after the onset of a stimulus. Such processes are considered to be automatic, not subject to strategic control, and preattentive, that is, they do not require conscious attention to be accomplished. In general, these early processes are thought to operate in parallel. Another set of processes occurs somewhat later in the information-processing sequence. These are considerably more under the subject's conscious control and require more attention. These two notions of automaticity and attention are of great importance in the carrying out of a choice RT task. Automaticity was discussed in greater detail in Chapter 3.

Blanco and Alvarez (1994: 78) point out that the intuitive notion that more intelligent subjects pay attention more efficiently than less intelligent subjects has been discussed in psychology for a long time. It is a generally accepted opinion in experimental psychology that general intelligence (commonly referred to as *g*) reflects not only individuals' knowledge of their world but also more basic skills for processing information that do not depend on the content of the information being processed (see eg. Jensen and Reed 1990, Neubauer 1990, Neubauer and Freudenthaler 1994). Isolating these basic skills and determining their releationship with traditional psychometric measures of intelligences has been the focus of a great deal of research in recent years (Blanco and Alvarez 1994: 77).

An issue in the use of reaction time measurement has usually been its imprecision. Zwitserlood (1989: 15), referring to Levelt (1978) and Cutler and Norris (1979) for similar distinctions, divides experimental methods into global and local tasks. Reaction time latencies obtained in grammaticality judgment tasks as used, for example, in this study, are always expressions of global reaction times. According to Zwitserlood (1989: 15), a global task measures the overall outcome of a number of processes. Such tasks include, among others, the cloze test, sentence verification, paraphrasing, and judgments on the grammaticality or semantic acceptablity of a sentence. Consequently, the responses in these tasks reflect the overall complexity of the stimulus material. Although global tasks can be timed, the responses cannot, as Zwitserlood points out, be unequivocally attributed to one level of processing or representation. For researchers who have been interested in the precise duration and location of the possible subcomponents of the whole sentence processing process, such imprecision has rendered this technique as insufficient. Attempts have been made to modify the task in various ways in an attempt to link the subjects' responses to a precise level of language processing, ie. to make them more local. Such tasks include naming, phoneme monitoring, word monitoring, and mispronunciation detection. For example, subjects have been asked to look for certain letters/sounds during the task. The location of the letters/sounds has then been varied. The value of this type of grapheme/phoneme monitoring task is based on the prediction that if letters/sounds placed at certain points in syntactic structure produced longer reaction times, this would be an indication of a reality of such a juncture in sentence parsing. The sequence and timing of

linguistic elements available to the subject have also been controlled for (Kilborn 1989). However, as Lehtonen (1984: 77) points out, the general aim of reaction time measurement combined with grammaticality judgment or other types of metalinguistic judgment has not been to discover the location or temporal duration of a particular linguistic process. Instead, reaction time studies are usually seen as an additional tool for the description of linguistic processes.

Reaction time measurement has been used as a tool in various kinds of linguistic tasks: First, it has been used in the language processing of individuals whose language competence has been seen to differ from that of normal native speakers of a language. Consequently, native speakers with impaired language processing ability such as aphasiacs or children with abnormal language development have been a favorite target group for researchers. The second language processing of nonnative speakers has also been studied, even if the number of such studies is not as great. Secondly, reaction time measurement has, of course, been used to shed light on language processing of native speakers. There has been a general shift in language processing research toward very precise local tasks, especially such techniques that measure the brain activity more directly, such as the measurement of event-related brain potentials of subjects exposed to various types of linguistic stimuli (see eg. Osterhout 1994).

7.2 Grammaticality judgments and reaction time measurement in L1 research

Wulfeck (1993) conducted a reaction time study of grammaticality judgments by children. She used grammaticality judgments and decision times for normally developing children of two age groups to determine the timing of parsing decisions and the influence of linguistic knowledge on parsing process in the processing of tape-recorded L1 sentences. Her results revealed, among other things, that both groups of children were better at detecting word order violations in a sentence than at recognizing errors of morphological selection (cf. Wulfeck and Bates 1991). Both groups of children also demonstrated very rapid integration of information during sentence processing. She concluded that language development seemed to involve more than the acquisition of linguistic knowledge, and that to become fluent speakers and listeners, it may be essential that children not only learn the important cues in their language but also process these cues rapidly and efficiently.

Blackwell, Bates and Fisher (1996) conducted an extensive study of the time course of grammaticality judgments by native speakers of English in three separate experiments, including a word-by-word grammaticality judgment task. The variables included the type and placement of errors. In contrast to L2 research, the items contained errors found in the speech of aphasiacs. Their results showed high correlations among the experiments involving both an incremental, word-by-word task and judgments of overall well-formedness. According to Blackwell, Bates and Fisher (1996), this suggested that the incremental tasks were tapping into the same decision-making process as is found online.. They conclude by stressing the importance of the careful construction of test items.

7.3 Grammaticality judgments and reaction time measurement in L2 research

Lehtonen and Sajavaara (1985) regarded reaction time measurement as a useful tool in psycholinguistic research as it has helped in isolating mental subsystems and various processual phenomena. The reaction-time method was applied to the testing of grammatical and pragmatic acceptability and to various test settings associated with lexical access and lexical decisions.

Lehtonen & Sajavaara (1983) used auditory tests of reaction time to compare second language learners' and native speakers' detection of ambiguity or acceptability of sentences. The subjects were Finnish learners of English as a second language and native English-speakers. The results showed that Finnish-speakers' reactions were generally similar to those of the native speakers. However, the Finnish-speakers were significantly slower in their judgments than the native speakers. Lehtonen and Sajavaara took this as an indication of a difference in processing strategies: native speakers appeared to have immediate access to meaning through context, while the nonnative speakers may have had to resort to problem-solving strategies. There were also clear differences between individual Finnish subjects, which was seen as a reflection of their English proficiency. In addition to the difference in judgment times, the native speakers were generally more unanimous in their decisions. For sentences ambiguous to natives, however, the Finnish subjects could usually see only one possible interpretation which appeared to be the more basic or common one. Lehtonen & Sajavaara (1983) acknowledge, however, that these findings do not necessarily reflect responses in real-life interactional situations.

According to Lehtonen & Sajavaara (1985), the basic assumption was that the rate and accuracy of the reactions reflected various processual features which determined the occurrence of L1-like constructs in L2 speech and writing. The grammatical acceptability judgment tests were assumed to touch upon the production mechanism in two ways: First, Lehtonen and Sajavaara assumed that if the listener accepts a pattern, it implies that the listener accepts the same construct also in his or her speech. Secondly, they assumed that a fast rate of decision-making indicated a high degree of availability of the data for the processing mechanism and a high level of automaticity in the access to the linguistic data base. In other words, the faster the reaction, the more automatic the choice. The first claim rested on the second claim: Lehtonen and Sajavaara (1985: 7) suggested that in the planning and production of utterances such constructs are chosen that are more readily available in memory, ie. are accessed faster. The first claim has remained untested since most of the research done in this area has sought to use reaction time measurement in connection of linguistic tasks involving receptive skills (see, however, Hedgcock 1992, Sutter and Johnson 1990); there is very little agreement even today whether the actual processes involved in language production and comprehension are the same, or even similar. On the other hand, the link between speed and automaticity is basic to psychometric studies in human information processing and the development of automatic skills. Consequently, the second assumption about the relationship between the access and use is plausible, for it rests on the existence of some kind of knowledge store (both declarative and procedural knowledge) which language reception and production may or may not share.

Tasks that involve the measurement of reaction time latencies typically have to do with the ease of processing, i.e. the learner's access and utilization of automatic and controlled processing (see eg. Robinson and Ha 1993, DeKeyser 1994, 1996). That is,

grammaticality judgement tasks combined with reaction time measurement can be used like any other task of language performance, with the added benefit of providing the researcher with the objective measure (and, indeed, operationalization) of language processing. As a matter of fact, it could be argued that grammaticality judgments used alone may turn out to be rather unreliable (see eg. Nagata 1987, 1989a, 1989b, 1989c) or, as Birdsong (1989a) points out, to measure an entirely new type of activities let in by the context of the test. Lehtonen and Sajavaara (1985:7-8) formulated three general hypotheses about the decision-rates in linguistic tasks:

(1) decisions at the boundary of two categories are slow (eg. ambiguous sentences);
(2) decisions based on conscious monitoring of explicit rules are slow;
(3) decisions containing active search in memory are characterized by slow rate.

Furthermore, they conluded on the basis of the results of the tests that the reaction times predicted the L2 processing mode in the way shown in Figure 13 below :

Reaction	Reaction rate	
	Fast	Slow
Correct	native-like categorical identification	ambiguous stimulus; application of conscious processing strategies
Erroneous	transfer on the automatic level of processing	

FIGURE 13 Reaction rate to various type of linguistic stimuli (from Lehtonen and Sajavaara 1985: 7-8).

The reaction-time measurement technique applied in the Jyväskylä Cross-Language Project was developed for the presentation of different types of linguistic stimuli. For example, Havia (1982) used tape-recorded L2 sentences, A. Lehtonen (1983) and Uusitalo (1988) visually-presented sentences. All these studies, like the present series of experiments, use acceptability judgments. Consequently, they are all based on the detection of error in the L2 sentences presented to L2 learners. In these studies, learners were not asked where the error was but rather, in general, whether there was an error. In this sense, reaction time measurement was used as a tool for research in language transfer. In other words, a situation was created where there was potential interference from an existing language knowledge system (either L1 or L2 or L3). The purpose was to see whether such interference would cause delay in judgments and whether it would lead to accurate judgments.

The notion of error detection has caused problems, however. As Birdsong (1989a: 107) points out, this in itself introduces a confounded variable for the use of metalinguistic judgment data to substantiate theory. Error detection is dependent on linguistic, perceptual and experiential factors. A major issue here is determining what subjects attend to when judging sentences.

Bley Vroman and Masterson (1989) used reaction time measurement as a supplement of grammaticality judgments in a sentence matching task. Kim (1993) studied the reaction times of Korean ESL speakers in a grammaticality judgment task. Sixty Korean speakers of English who had lived in the United States were tested for their ability to judge the grammaticality of English sentences across different grammatical categories. Kim found that those nonnative speakers who had learned English at an earlier age performed faster in the task performance. Kim interpreted the findings to suggest that a sensitive period exists in second language acquisition.

Similarly, Shim (1993) found in her study of Korean ESL speakers that the mean reaction times for native speakers were significantly faster than those even for early bilinguals. Again, she took these findings to indicate that the ultimate proficiency in a second language was affected by the age at which learners began learning the target language. Robinson and Ha (1993) used reaction measurement and grammaticality judgments to find support for Logan's instance theory (see Chapter 3).

Robinson (1994) investigated the acquisition of noun incorporation processes by adult learners of Samoan L2 while comparing it to the performance of native speakers of Samoan. The method used combined grammaticality judgments with reaction time measurement and response certainty measures. More specifically, Robinson was interested in the influence of a proposed implicational hierarchy and constraints of Universal Grammar on the acquisition process.

Juffs and Harrington (1995) investigated subject extraction in English by native speakers of Chinese who were advanced learners of English. They used grammaticality judgments of full sentences and in the word-by-word condition. The results indicated that the difficulty that the learners had with subject extraction sentences was due to parsing and not grammatical competence.

7.4 The general research question

As can be seen from the short review above, grammaticality judgments, with or without reaction time measurement are widely used in second language research. In the following chapters, the results of three RT judgment experiments will be discussed, all using material based on errors found in L2 learners' productions. In this respect, the experiments will differ from the majority of the studies described above (cf. Hedgcock 1993). The RT judgments have been employed in these experiments to investigate learners' reactions in cases where their interlanguage knowledge may still be indeterminate and show the effect of declarative and/or procedural knowledge of other language systems, ie. language transfer. The specific questions asked concern the nature of the data obtained by RT judgments, its relationship to a number of cognitive and affective variables, and L2 proficiency in general.

8 Experiment 1

8.1 Background: language transfer and L3 learning

The purpose of Experiment 1 was two-fold: first, to investigate the influence of learners' L2, English, on their weaker L3, German; and second, to discover whether learner performance on English L2 sentences depended on learners' L1.

According to Singh and Carroll (1979: 51), there is no *a priori* reason to think that L3 learning is any different from L2 learning. The only difference between the two phenomena is the amount of previous language experience. However, some studies on the role of transfer in L3 learning suggest that there might be a difference between the processes involved, even if it is hard to tell how far-reaching these differences really are, that is, whether the processes in actual fact are radically different from each other.

Accordingly, LoCoco (1976: 59) is able to report on the basis of a study that bilinguals learning an L3 seemed to make proportionally fewer interlingual errors. This appears to be due to the increased experience with languages as the learner is then more aware of the fact that reliance on one system may cause interference in the other system. Metalinguistic awareness may also play a role (see eg. Bialystok 1988).

There also seem to be differences between native language influence (L1 influence) and non-native-language influence (LNN influence) in the areas where transfer manifests itself in the learner language. Ringbom (1985b: 39) points out that LNN-influence occurs frequently at the level of lexis, whereas it is much less significant, if it exists at all, in grammar and phonology. Studies in this field can be extremely meticulous including careful classification of lexical items according to their similarity or dissimilarity relative to another language (see eg. Ard and Homburg 1983). By contrast, L1 influence most frequently appears on the phonological level of language. Interestingly, the role of lexis as opposed to grammatical rules in language learning may be of a much greater significance than previously thought. Widdowson (1989), for example, suggests that lexis might play a greater role in language learning. Indeed, if one accepts his view of the development of communicative competence, rules of language are not generative, but their application is based on creative adaptation of formulaic frameworks, patterns and a kit of rules that have been developed from patterns of lexical co-occurrence, lexical clusters, whose meaning is apparent (Widdowson 1989: 136). In a generative view of grammar, rules such as a suffixation rule that adds the past tense marker at the end of regular English verbs (*walk-walked*) have a major role in language processing and learning. Logan's (1988) instance theory of skill and automaticity in language processing rests on the notion of instant retrieval from the memory rather than ever-improving use of algorithms, ie. 'rules'. However, Ringbom (1986: 159) does not exclude the possibility that LNN influence on other areas than lexis may be more frequent if the LNN knowledge is sufficiently automatized.

Ringbom (1985b: 56-57) further states that the difference between L1- and LNN-influence lies in the fact that LNN-influence tends to surface in borrowing only, ie. in contexts where the search for a lexical item has activated another item, usually formally similar, in a non-native language, whereas lexical influence from L1 can manifest itself in either borrowings or lexical transfer. In reception, this would call forth wrong identification (cf. Faerch and Kasper 1987).

According to Ringbom (1985b: 39), the studies in LNN-transfer have concentrated either on the influence of two European languages (one of them L1 of the subjects) on a third European language, or they have investigated the LNN-influence from another European language on Asian or African learners of English or French, often including comparisons with L1-transfer (for the latter type, see eg. Ahukanna et al. 1981). The present study thus falls in with the former type of studies with the additional element of the factor of perceived distance (eg. Kellerman 1977, 1983) playing a part; L1 influence on English as a foreign language was also under study as the subjects included language learners with widely different L1s (namely Finnish and German). A great deal of research has also been done on the differences between Finns and Swedish-speaking Finns in the learning of English as a foreign language. Sjöholm (1979: 93) noted how many of the errors among Finnish-speaking learners of English seemed to be due to Swedish influence whereas among Swedish-speaking learners of English the influence from Finnish, ie. from L3, was extremely rare (see also Ringbom and Palmberg (eds.) 1976). Ringbom (1987) has studied the role of the first language in foreign language learning in Finland with the special regard to the differences that exist between the Finnish-speaking and Swedish-speaking Finns in the learning and processing of English.

In addition to these more or less linguistic considerations, it is probable that LNN transfer is influenced by factors similar to L1 transfer (Ringbom 1986: 159). Thus, the language learner's age, stage of learning, mode of learning, personality, and sociolinguistic context will certainly play a role in the learner's willingness to transfer.

In Experiment 1, an attempt was made to investigate non-native language influence on other areas in addition to lexis, including even items concerning orthography and bound morphemes. The research focused on discovering what kind of linguistic phenomena the learners tended to transfer, and furthermore, what kind of role the learners' L1 played, ie. whether there were any differences between the speakers of Finnish and German in their judgments of English. During the past ten years, the importance of lexical items both in language processing and language learning has gradually been given the attention it deserves, although the difference between system-learning and item-learning has been known for quite some time (see eg. Ellis 1994). According to this view, the language learner cannot start learning L2-systems such as phonology, intonation, morphology, syntax and semantics directly but he first has to know at least a number of items. In L2-learning, in contrast to L1 learning, this means that the beginning learner constantly assumes equivalence between what is new, ie. L2-items, and what is known from before, ie. L1-items (Ringbom 1985a: 11; cf. Cruttenden 1981).

One of the factors determining how easily such cross-linguistic equivalence between L1 and L2 items can be established is perceived distance between L1 and L2 (Kellerman 1977). On the quite concrete level this means that, for example, the existence of cognates has a facilitating effect in foreign language learning, especially as far as receptive competence of the language learner is concerned (cf. Odlin 1989).

Within this context, lexical transfer can be seen as application of a learning hypothesis: that lexical items are either translation equivalents to, or have the same semantic features as, items in the learner's L1 or some other language he knows well (Ringbom 1983a: 210). The latter point is of especial significance within the framework of this study, especially as Ringbom (1985b: 41) notes that the cross-linguistic influence between non-native languages in a European context has been shown to occur in the first place in lexis. In other words, it was hypothesized that Finnish subjects would show an inclination to accept lexical items

or idiomatic strcutures based on English in the German and Swedish sentences presented to them.

Finally, a short note on terminology used may be in place here. Some researchers make a distinction between lexical and semantic transfer. Stedje (1977: 145-147) uses both categories of lexical and semantic "interference" in her study on the influence of Swedish on the German spoken by Finnish students in Sweden. She uses the term lexical transfer when speaking of the language learners' use of Swedish lexical items (either direct borrowings or lexicalizations) in their German; the term semantic transfer is used to refer to the attribution of L1 or L2 semantic content on a L3 (German) lexical item. It is to be noted that 'false friends' belong to this semantic 'subcategory'. However, Stedje (1977: 147) points out that it is not always possible to differentiate between the phenomena. Lexical transfer seems to be the general term in use, however (eg. Ringbom combines both lexical and what is understood as semantic transfer under the same heading; Juhasz (1973: 458), for example, on whom Stedje seems to rely to a certain extent in her discussion of transfer, prefers the term semantic interference.

8.2 Subjects

Originally, three groups of subjects were used in these tests as they were carried out in Finland. A fourth group was added on the basis of tests run in Kassel, West Germany. Group 1 consisted of 20 L1 Finnish university students of English; they will be referred to as the Finnish students or *Finns*. They all studied English at the University of Jyväskylä at an advanced level (either at a cum laude or laudatur level). On the basis of earlier studies (see eg. Havia 1982: 50) it was assumed that there was no difference between cum laude and laudatur students with regard to their performance in the tests in general. All the Finns had spent the required period of time in an English-speaking country (two months) and eleven had resided more than six-months in such a country.

In analyzing the results, the Finns were additionally divided into two subgroups according to whether they studied German at the University of Jyväskylä; all Finns had studied German at least for three years at school. These two subgroups, Groups 1a and 1b, each contained ten subjects and will be referred to as *students* and *non-students of German*. As the number of Finnish subjects who studied Swedish at the university level was small, only five, it was not thought to be productive to analyse the results with the subjects' proficiency in Swedish in mind, although a note of this difference will be made if need arises. The length and level of language studies was thus used to control the subjects' level of language proficiency. This approach was not adopted without reservations; however, it was felt to be the simplest way of dealing with subjects with such varied background in language studies.

Group 2 included fifteen native speakers of English, nine of whom were Americans, five Englishmen, and one Australian. All Englishmen were teachers of English either at the university or private language schools. Only one of the Americans was a teacher of English, one was a visiting lecturer (a linguist), and the rest were either university or high school students. This group will be referred to as *native speakers (NSs) of English.*

Group 3 consisted of thirteen native speakers of German. All of them were or had been teachers of German in Finland. Only one of the subjects had resided for longer than a year in an English-speaking country, and only one had studied English at the university

level. Consequently, this group could be characterized as being 'heterogeneous' as far as their English proficiency was concerned. Most of the Germans in this group had studied English at school for 8-10 years, one had studied English later, and two had studied English at school for about five years. This group will be referred to as *German teachers*.

Group 4 included 20 German students of English. Thirteen of them had spent the required time ("Auslandssemester") in an English-speaking country. They will be referred to as *German students*, or German students of English.

8.3 Design of the materials

Test items were designed to reflect both lexical and syntactic transfer. Thus, the material of these tests consisted of sentences including phenomena from all levels of language. In choosing the materials, the focus of interest lay in the structures that were assumed to be potential objects of transfer. Valuable insights into the nature of errors among Finnish students of English (both beginning and advanced) were gained through four error analysis studies done at the Department of English in Jyväskylä (Korhonen 1983; Lamminpää 1980; Taskinen 1981; Herranen 1977). Two studies on typical errors of English/German cross-linguistic influence also provided some information (see Morrissey 1983; Grauberg 1971) for the design of the test items.

The test items were grouped into seven blocks according to the language and task type. In addition to subjects' L2, English, the experiment included items in German and Swedish, ie subjects' L3 and L4. The two task types used in the experiment were the judgment of grammatical acceptability of isolated grammatical and ungrammatical sentences and the judgment of correct and incorrect translation of English L2 sentences into subjects' L1, L3 and L4. Block 1 consisted of English sentences, Block 3 Finnish, Block 5 German and Block 7 Swedish sentences in isolation. Block 2 consisted of English/Finnish translation pairs, Block 4 English/German and Block 6 English/Swedish translation pairs. All blocks contained items that had structural equivalents in all the other blocks (only exceptionally, however, in blocks involving Finnish).

The test sentences contained instances of both grammatical and ungrammatical use of a particular linguistic structure. The length of the sentences was strictly limited; as the sentences were not to be longer than one line, it was decided, after some experimentation, that in order for the sentences to be fully visible on screen, they should not be longer than 29 characters (blank spaces included) (10 cpi). This put, naturally, heavy constraints on the material chosen. The test items are usually of the type SVO, SVA, or sometimes SVOA and SVAA; there were some instances of embedded structures, mainly very simple infinitive clauses, or adjective attributes. In addition, all task blocks contained sentences of maximum length, some of which were patently ungrammatical, others clearly grammatical. The original idea was to use these sentences to provide a baseline for the judgment times of patently grammatical and ungrammatical sentences; as it turned out, however, these sentences proved to be more problematic or ambiguous than the test sentences proper, and they were never used for this purpose.

8.3.1 Grammaticality judgment tasks

In the experiment, there were four different task groups consisting of sentences in isolation. Block 1 consisted of 91 English sentences (items 101-191), of which 43 were grammatical and 48 ungrammatical (see Appendix 1). Block 5 consisted of 34 German sentences (items 501-534). Again, the items included both grammatical and incorrect use of particular linguistic structures. Similarly, Block 7 contained 29 Swedish sentences (items 701-729) representing grammatical and incorrect use of linguistic structures. At a somewhat late stage, it was decided to include some Finnish sentences in the test, more in an experimental sense. Thus, Block 3 contained eight Finnish sentences (items 301-308), three of which were ungrammatical, two grammatical, and the remaining three ambiguous. The ambiguous sentences were later excluded from the statistical analysis; thus, the final number of Finnish sentences was four. These blocks will be referred to as English (Block 1), Finnish (Block 3), German (Block 5) and Swedish (Block 7) blocks.

8.3.2 Translation judgment tasks

Yet another approach to transfer is provided by translation science. Translation theory and contrastive analysis are closely connected (see eg. James 1980: 4). Both are interested in cross-linguistic communication; James includes them both in the same domain of interlingual linguistics (although he speaks of 'translation theory'). The notion of transfer is also important for translating "as a cross-linguistic type of communicative mode of production" (Toury 1986: 80). Transfer in translation, however, is not to be understood to operate in the same way as it does in other modes of language production. First, the role of the source language (SL) is limited to the very utterance to be translated, that is, it is not retrieved from the speaker's knowledge of the SL, ie. system as such but, instead, it is given. Secondly, the translator is dependent on the SL utterance in the sense that he must try to preserve certain features invariant, at one level or more; in other words, he must try to preserve something that has already been encoded in language. There is an additional role for the translation process: it can be hypothesized that language learners use it as a cognitive strategy during language processing. Beginning learners especially may first, either unconsciously or consciously, translate an item into their L1 (cf. Tomasello and Herron's (1989) garden-path technique). There may even be, however, as Klein (1986: 13) points out, a tendency for experienced speakers of a second language to rely on comprehension strategies that have their origin in the use of L1 (cf. Bates *et al.* 1982).

A quite interesting approach to the interrelationship between transfer and translation is provided by Danchev (1982: 49). According to him, if translation consists mainly of interlingual transformation(s), transfer can then be considered isomorphous translation, where the obligatory transformations included in the process have not been carried out. Negative transfer is thus a result of partial translation, which distorts the structure of the target language and leads to erroneous utterances. Moreover, where such errors occur, they are the result of insufficient monitoring of the respective utterance by the translator. This view is quite consistent with the fact that good translation is a task that requires quite specific metalinguistic skills and certainly a certain amount of conscious control.

As it has already become apparent, transfer is especially likely to occur in connection with translation tasks. In the present study, such translation tasks consist of a short English sentence in isolation and its - correct or erroneous - translation into another language. The highest level of such segmentations is thus the level of sentence. The starting point for

translation is a sentence in English, and we can now hypothezise that the subject as translator uses only the exact structures embedded within sentences. The task that they were asked to carry out implied that, first, the subjects made their own translation of the source sentence, and then compared the result with the target sentence, which may or may not have been in their native language. (However, the results imply that some other strategies may have been in use.)

Three task groups were included in the experiment: Block 2 consisted of seven English/Finnish pairs (201-207), some of which were ambiguous, others clearly correct or incorrect. This block was designed for finding out the Finnish learners' reactions to test items containing sentences in their native language. Block 4 consisted of 23 English/German sentence pairs (401-423). The majority of the items contained similar grammatical problems as the sentences in isolation. Block 6 consisted of 19 English/Swedish pairs (601-619), which, again, dealt with corresponding linguistic categories as the other language blocks, particularly Block 7. These blocks will be referred to as English/Finnish (Block 2), English/German (Block 4), and English/Swedish blocks (Block 6).

To help the subjects to focus their attention to the non-English sentence in the translation judgment tasks, the lower half, that is, the background of this item was covered by pale yellow transparent sticker tape. The translation pair tasks looked like the following:

411 She saw Helen come

Sie sah Helen kommen

8.4 Research hypotheses

The following hypotheses were made concerning the learner vs. native speaker performance and the learners' acceptance of erroneous sentences.

Hypothesis 1. As regards the native speaker vs. language learner responses, it was hypothesized that the language learners would need more time for the processing of the linguistic input than the native speakers, and that the processing time would be reflected in the RT latencies as the native speakers' decisions on the grammatical acceptability of the sentences would be largely automatized (Lehtonen and Sajavaara 1983: 103).

Hypothesis 2. Native speakers were expected to show greater accuracy in their grammaticality judgments and the judgments of Finnish university students of English and presumably those of the German subjects, would vary considerably.

Hypothesis 3. With regard to the three learner groups, it was hypothesised that some kind of difference could be found between Finns and Germans in English L2 reaction times. The Finns have been noted to be at a disadvantage as far as perception of speech was concerned. No similar problems had been reported about the production and perception of written language (Lehtonen 1983: 6). The findings of Havia (1982: 113) also seemed to support this view: it was easier for the Finnish subjects to judge sentences when they were presented visually. However, as Lehtonen (1983: 15-16) points out, it seems reasonable to assume that a Finnish language learner will also have problems of a quite specific nature in processing English. Such problems have their origin in the writing system of Finnish, which is usually regarded as almost fully phonological. This affects the organisation of, and access to, lexical memory. The reliance on orthographic form is also reinforced in language

teaching situations in which the written medium is often used (see also Ringbom 1987). Thus, Ringbom's (1982, 1987) findings about Finns' difficulties in the processing and learning English, if contrasted with their Swedish-speaking peers, would support a hypothesis according to which Finns would very likely be slower than the German students of English in Group 4. Moreover, this is supported by a general assumption that language learning is easier between the speakers of typologically similar languages. However, it was not clear how their performance would be related to that of the subjects in Group 3, who were university teachers of German and whose English proficiency was generally speaking not as good. It was tentatively assumed, though, that the Finns would be faster and more accurate than the teachers.

Hypothesis 4. It can be assumed that differences in the processes of information production and perception between two languages are also reflected on the interlanguage processes; the language learner is likely to apply the native language processing mode when processing and understanding messages in the non-native languge (see Sajavaara and Lehtonen 1986: 1444; Bates and MacWhinney 1989, Kilborn 1989). What is said here about the native language surely applies to L2 if L2 only is sufficiently automatized. As to cross-linguistic influence between two non-native languages, it was assumed that the Finnish language learners' linguistic knowledge of English, as university (ie., advanced) level students, would be at least in part internalized and automatized to the effect that transfer from it could take place. It was hypothesised that the Finns would show a tendency to transfer linguistic features of their 'strong' L2 to another non-native language, whether these features were grammatical rules, that is, relating to syntax and morphology, or semantic components. The focus of interest was the influence of English on German although the tests contained also items in Swedish.

Hypothesis 4a. The influence of English was expected to show on the Finnish subjects' judgments on German in that they would show a tendency to accept German structures modelled on English.

Hypothesis 4b. Secondly, the influence was expected to appear in their RT latencies in that a fast acceptance of deviant L3 sentences based on English L2 would be taken as an indication of the use of routines developed for L2 processing.

Hypothesis 4c. As to the translation pairs, it was hypothesized that the subjects would show a greater tendency to accept structures in L3 and L4 that were based on English L2 in translation pairs than in isolated sentences.

8.5 Procedure

The sentences were typed, and then transferred onto transparencies, which were then attached to frames. Likewise, the test instructions in Finnish, English and German were typed and transferred to transparencies. It is to be noted that in this experiment, the subjects were not told to judge on grammaticality or acceptability of the sentences in question but they were simply required to decide whether the sentence is in their opinion correct or not. It was hoped that such instructions would produce as natural and fast a reaction as possible. As the sentences contained errors that can be traced to specific rules of grammar, the term grammaticality judgment can be used (or acceptability; the nature of the material is such that the correct sentences are both grammatical and acceptable, and incorrect sentences both ungrammatical and unacceptable, with some exceptions). The use of the term 'correct' may

have led some subjects to think of school grammar and formal rules of grammar; yet, the example shown to the subjects was a sentence containing a semantic anomaly, a phenomenon which was not regarded by linguists to belong to the area of grammar proper.

The actual testing took place in an experimental laboratory at the Department of Communication, University of Jyväskylä. Before the beginning of the test, subjects were 'eased' into the test situation by giving them the same instructions orally as they were later shown on the screen. Also, the subjects were asked some background information about themselves concerning the length of their stay in an English-speaking country, the year of their studies, the number of languages studied either at school or at the university, and the grades in the school report.

The subjects were in a separate sound-proof room. They had a screen in front of them, on which the sentences were projected by means of a tachistoscopic slide projector. The projector was connected with an electronic RT measuring instrument. The time of exposure was 2100-2150 ms, and a light transistor started the clock at the beginning of each exposure. On the table in front of the subject, there was a panel of the buttons for "yes" and "no" reactions. An arbitrary cutoff value of 9999 msec was used to exclude prolonged RTs. The subjects were asked to push the button marked with green as soon as they had decided that the sentence was correct; if they thought it was not correct, they were to press the button marked with red. Approximately 700-800 msec before the presentation of each sentence, the subjects heard a short signal (roughly 700 msec in duration). This was done to produce a priming effect whose purpose was to direct the subjects' attention to the screen (since the experimenter controlled the presentation of the slides, the time gap between the sentences was not identical in every case).

The instructions were shown on the screen. Each subject group received the instructions in their native language. The subjects saw the task groups that involved the same language always in the same order: the translation pair task always preceded the corresponding task group with isolated sentences. Thus, the order of presentation was: English, English/Finnish, Finnish, English/German, German, English/ Swedish, and Swedish. About half of the Finns saw the Swedish task groups before the ones involving German. Before each task group, the subjects were shown a new set of instructions.

The order of the task items was changed during the tests so as to avoid the effects of practice and potentially prolonged RTs that the first sentences would otherwise receive. The erroneous items in the blocks consisting of sentences in isolation were presented first in order to prevent learning from taking place; a few sentences were shown in reverse order so as to serve as distractors of a kind: it was thought possible that some subjects would notice the strategy behind the arrangement of the items.

8.6 Statistical analysis

The following methods were used in the statistical analysis of the results: means, standard deviations, t-tests, correlations, and ANOVAs. The latter were applied to measure the statistical significance of the differences between the mean RTs. All calculations were carried out with the help of the SPSS program package on the Jyväskylä University mainframe. The guidance service of the University Computer Centre was used in devising the data and system files and in the statistical analysis.

8.7 Results and discussion

8.7.1 Hypotheses 1 - 2

The hypothesis that the NSs of English would have lower reaction times than the language learners and that they would also be more accurate in their acceptability judgments was borne out by the results. Tables 1 and 2 below list the mean RTs and scores across the NSs and the three learner groups. The score was calculated by giving the subjects one point for every 'correct' judgment; note that even the slightly ambiguous sentences were scored.

The ANOVA revealed that the difference between the mean RTs was extremely significant. A post hoc Scheffé procedure was used to discover which groups were significantly different from each other (at the .05 level). The Finns and the German teachers were significantly slower than both the NSs of English and the German students, and while the Finns were faster in their judgments than the rather heteregeneous group of German teachers, the difference was not statistically significant. The German university students of English performed the tasks in English much quicker than their Finnish peers. In fact, their mean RT comes very close to the mean RT of the native speakers as the statistical difference between their mean RTS and those of the NSs of English was not significant. The differences between the NSs and the Finns and German teachers, respectively, were significant. Table 3 below shows the source of variance, the F-value and its significance for

TABLE 1 The mean RTs and standard deviations (in msec) for different language blocks.

	Finns (N=20)	NSs of English (N=15)	German teachers (N=13)	German students (N=20)
Finnish (k=4)	2389 (684)	n.a.	n.a.	n.a.
English (k=91)	2870 (496)	1998 (377)	3400 (909)	2178 (523)
German (k=34)	3379 (836)	n.a.	2382 (386)	1785 (340)
Swedish (k=29)	3082 (724)	n.a.	n.a.	n.a.
English / Finnish (k=4)	3263 (526)	n.a.	n.a.	n.a.
English / German (k=23)	3962 (800)	n.a.	3473 (541)	2306 (506)
English / Swedish (k=19)	3784 (790)	n.a.	n.a.	n.a.

TABLE 2 The mean scores and standard deviations for different language blocks.

	Finns (N=20)	NSs of English (N=15)	German teachers (N=13)	German students (N=20)
English (k=91)	74.7 (4.59)	81.7 (4.32)	67.85 (8.27)	65.7 (7.36)
Finnish (k=4)	3.9 (.49)	n.a.	n.a.	n.a.
German (k=34)	23.1 (4.11)	n.a.	31.5 (1.5)	29.4 (2.54)
Swedish (k=29)	22.1 (2.21)	n.a.	n.a.	n.a.
Finnish / English (k=4)	3.7 (.57)	n.a.	n.a.	n.a.
German / English (k=23)	13.9 (2.69)	n.a.	17.8 (1.69)	14.4 (3.25)
Swedish / English (k=19)	12.4 (1.79)	n.a.	n.a.	n.a.

the mean RTs of the English sentences across all subject groups. The ANOVA also showed that the difference in the scores was statistically significant. A post hoc Scheffé procedure revealed that in this, the NSs were significantly more accurate than the three learner groups. The Finns, on the other hand, were significantly more accurate than the two German groups. Significantly, the German students' judgments were less accurate than those of the Finnish learners, and their mean raw score was actually a little lower than that of the German teachers, who did not even study English.

TABLE 3 Source of variance, F-value and its significance for the mean RTs and scores for Block 1 (English sentences) across the four subject groups.

Block 1 (English) (k=91)	Source	Degrees of freedom	Means square	F	p
Mean RT	Between groups Within groups	3 64	6242139.10 340575.58	18.32	.000
Mean score	Between groups Within groups	3 64	851.82 39.25	21.70	.000

TABLE 4 Sources of variance, F-values and their significance for the mean RTs and scores for English/German and German sentences for the Finns, German teachers and German students of English.

	Source	Degrees of freedom	Means square	F	p
German (k=34)					
Mean RT	Between groups Within groups	2 50	12892598.86 345887.87	37.27	.000
Mean score	Between groups Within groups	2 50	336.13 9.42	35.68	.000
English / German (k=23)					
Mean RT	Between groups Within groups	2 50	14273346.32 410729.28	34.75	.000
Mean score	Between groups Within groups	2 50	93.04 7.44	12.51	.000

As to the differences between the Finns and the Germans in the German-language tasks, Table 4 shows the source of variance, F-values and probability in the German-language blocks. A *post hoc* Scheffé procedure revealed that the Finns were significantly slower and less accurate (at .05 level) in judging isolated German sentences than both German L1 groups. The Finns were also significantly slower than the German students (but not teachers) in judging translation pairs. Interestingly, the German students were significantly faster than the teachers in carrying out both types of tasks in German. On the other hand, the teachers were significantly more accurate in their judgments of the translation pairs but while their raw mean score was slightly higher than that of the students for isolated German sentences, the difference was not significant. This seems to suggest that the inaccuracy of the German students' judgments in English and in German / English translation pairs may not necessarily have been due to the speed of their reactions only or be a reflection of more or less conscious answering strategy, since they seemed to be able to judge sentences in their L1 practically quite as accurately as the German teachers.

To sum up, it was assumed that the native speakers' judgments would be faster than those of the learners. The results confirmed this hypothesis as to the Finnish learners, both in the case of English and German. However, the German students of English turned out to be almost as fast as the native speakers of English in the English-language task, and no significant differences emerged between these two subject groups as to the speed of their judgments.

It was also hypothesized that the native speakers of English would be more accurate in their judgments than the learners. This was confirmed by the results, as shown in Table 2 above. However, it is worth noting that the native speakers were not unanimous in their judgments; the judgments did not uniformly match the expectations, either. This seems to be due to a number of sentences on whose grammaticality the native speakers did not seem find agreement.. Most of these sentences had to do with the use of the articles (138, 140, 141), an aspect of the grammatical system notoriously difficult to deal without context;

some had to do with the use of the correct past tense form (148 and 149) or the correct infinitive form after the verb *go* (170 and 171), both items in which there seems to be a great deal of variation among English speakers. Perhaps not surprisingly, the purely conventional matter of correct spelling seemed to be a cause of disagreement (184 and 186).

Thirdly, it was hypothesized that the Finns would presumably be faster and more accurate than the German teachers, whose English proficiency was, after all, not assumed to be as high as that of the Finns. However, no such clear hypotheses were made as to the relationship between the Finnish and German students of English. Even so, it was tentatively assumed that the Finns would be slower than their German counterparts. This was, overall, supported by the findings. The difference in the accuracy of judgments was quite unexpected, however.

The results could be explained as follows: First, the German students' faster RTs could be a function of their mother tongue. Because of the typological similarity between the languages, it will be easier for the Germans to gain a certain surface structure familiarity with English from the very beginning, ie. the starting point for the Finns and Germans is different (cf. Ringbom 1987). If this is the case, shorter RTs could also be expected from the other German group, as well. As a matter of fact, considering the level of their English proficiency, the German teachers performed relatively well, even if the Finns were faster, which was to be expected on the basis of the difference in language proficiency. The difference between their means was only almost significant, which is not at all what was to be expected. All this seems to point to native speakers of German having a considerable advantage over Finns in both the learning and the processing of English, which manifests itself in the greater speed (near native-like) in the making of acceptability judgments.

The relative inaccuracy of the German students was equally striking. It seems to suggest that the students resorted to automatic processing strategies in which they were able to utilize procedural knowledge from their L1 or interlanguage developed on the basis of hypotheses about the similarity of German and English. Consequently, they were able to respond quite fast. Another explanation also offers itself: The Germans' responses to German sentences suggest that there is a difference between the two groups of Germans: the German teachers were more careful and accurate in their judgments. This could have been a function of their profession, which made them more aware of linguistic form. The German students of English had probably learned their English in an environment that favoured a more communicative approach to language learning and focused less on linguistic form although, of course, nothing certain is known of the mode of the instruction that the German subjects had received. It can be assumed though that the 'teacher' group had mostly been taught with the traditional grammar-translation method; at least it is reasonable to assume that their present occupation was heavily influencing their judgments forcing them to concentrate on form. There is no data available concerning the background of the German students (except their length of residence in an English-speaking country) but it seems that during the seventies certain states in the Federal Republic of Germany moved over to the communicative competence approach in language teaching (see Jung 1981). Thus, it may be that the difference in language teaching methodology has influenced the subjects' attitude towards language, hence their faster but inaccurate judgments.

The latter factor could also partly explain the relatively slow performance of the Finns if compared with both the NSs and the German students: language teaching in Finland, at least for the age group represented by the Finnish subjects, has very likely followed the more traditional routes, if not entirely at least in most part.

There may be other factors at play as well: As Lehtonen (1983: 15-16) points out, a Finnish language learner may have processing problems caused by the writing system of Finnish which is usually regarded as almost fully phonological. According to Lehtonen (1983: 16), it could be that the writing system of Finnish may affect the way Finnish language learners store words of a foreign language, and then further on processes it. This influence shows up in the difficulties Finns have in the production and perception of speech (Lehtonen 1983: 6). This view presupposes that instead of there being two distinct files in the lexical memory, ie. phonological and orthographic, only one form would be needed in Finnish which allows for both visual and phonetic access (Lehtonen 1983: 16; cf. Levelt 1989). The reliance on the orthographic form is also reinforced in the language teaching situation in which the written medium is often used. Thus, the results may also have been affected by some very fundamental differences in the processing of the written text by the Finns.

8.7.2 Hypothesis 4

It was initially hypothesized that the Finns would show a tendency to accept structures in their L3 that were based on L2. The analysis of individual sentences of this type would seem to support this notion. These sentences have been discussed in detail elsewhere (Alanen 1987, 1991); therefore, only the most interesting cases will be described below with the aim of demonstrating the very real differences that this method appears to reveal in the subjects' performance.

It was hypothesized that Finnish learners of English would show a tendency to accept structures based on English in their 'weaker' foreign language. The focus of the investigation lay in discovering how the language learners would judge sentences in German. As regards the hypothesis made earlier on the nature of language transfer, it was assumed that all the Finns had internalized and automatized the corresponding English structures to the extent that transfer was possible. Above, it was discussed how transfer might operate within different modes of information processing and types of knowledge. The reaction times were considered to reflect the degree of automatization of a particular linguistic structure. If the language learner quickly accepts a deviant structure containing an error based on L1 or any other language with sufficiently automatized access to relevant knowledge structures, it could be assumed that she has been assisted in the decision-making process by the existence of such ready-made subroutines. Following Lehtonen and Sajavaara (1985), those L2-based sentences that were accepted fast were held as prime evidence of transfer of L2-based automatized processes. To rate which sentences were judged quickly, the sentences were ranked according to the mean RT they had received. These mean RTs were then compared with the mean for the whole language block. Those sentences that were judged faster or at the same speed as the mean RT for the whole block were considered candidates for language transfer. Native speaker judgments of the Germans were also used for comparison but since they were generally much faster than the Finns, no direct use was made of them for the ranking. Of course, this only demonstrates how non-native like even the fastest RT judgments of the Finns were.

However, it turned out that fast acceptance of L2-based structures in L3 occurred mainly in Blocks 4 and Block 6, that is, during the translation pair tasks. This further supports the tentative view that in such a translation judgment task there are many intervening factors, one of them being the "speaker" type of the subject. In such a task a quick type is favoured. Hypothesis 4c, ie. that transfer would take place more readily in

connection with translation, would also find some support. For isolated sentences containing the same structures, the RTs tended to be relatively long. It is to be noted, though, that the learners, on the whole, showed a tendency to accept these structures even in isolated sentences. The longer RTs seem to indicate that they had to search their long term memory and resort to less automatized processes to find the solutions.

In the following discussion, translation equivalence was used as a basis for contrasting structures. In doing so, many problems connected with this notion were overlooked (cf. Sajavaara 1981b: 41; Odlin 1989: 46-47). Similarly, although for example case grammar would have provided a perhaps more insightful approach to the study of transfer, the analysis was nevertheless kept quite near the surface structure level. The results of the tests will be discussed on the basis of the language subsystems on which the erroneous 'products' occurred.

Items that tended to be accepted fast included item 407:

*407 He was given a book.
Er wurde ein Buch gegeben

It was accepted fast by 60% of the Finns. This is a case of accepting the passive formation based on L2 in L3. The passive is one of the areas in which the Germanic languages under discussion, while sharing the underlying principle of the passive formation, also differ from each other, whereas the passive transformation such as it occurs in English, German and Swedish is not possible in Finnish. The Finnish passive indicates that the action of the verb is performed by an unspecified person, ie. that the agent is impersonal. There is no grammatical surface subject (Hakulinen and Karlsson 1979: 255). It roughly corresponds to English *one* and German and Swedish *man* (Karlsson 1983: 146). The English passive has in addition a feature which German does not share; in German, it is not possible to transform the indirect object of the active sentence into the subject of the passive sentence. Consequently, it is possible to study the subjects' judgments on task items that do not have direct structural equivalents in L1 or L3s. It was assumed that the Finns would show a tendency to judge those German sentences to be correct where the passive is formed acording to the English pattern. This tendency was thought to be strengthened by the metalinguistic knowledge of the language learners: there is a category traditionally known as 'passive', and this probably was presumed to affect the learner's perception of it.

The Germans' responses to the task items involving the translation of English L2 into their L1 were quite fast. This time 80% of the German students rejected the incorrect translation. The German teachers were even more accurate: 12 out of the 13 rejected the translation. Even the mean RTs were quite similar, 2481 msec for the teachers and 2225 msec for the students, which was exceptional, as the latter were usually faster in their judgments. The Finns, who had to judge between two foreign languages, were again noticeably slower, with the mean RT of 3604 msec.

When the same construction occurred in isolated sentences in German, the tendency evident in item 407 became even clearer. Whereas the native speaker judgments for sentence 512 **Ich wurde ein Buch gegeben* were quite undivided, all the NS subjects judging the sentence to be incorrect, the Finns showed a great deal of variation: 55% of the Finns accepted the sentence, five being students of German. (The NS responses to sentence 511 *Mir wurde ein Auto gegeben* were slightly more divided, which could have been due to the content of the sentence.) The Finns, again, showed a great deal of variation as a total of 50% of the Finns found this sentence correct.

Again, the Finns' responses reveal a large degree of uncertainty even among the university students of German about the correct passive construction in German. There seems to be a tendency to accept a structure which was directly modeled on the corresponding English structure. This appears to be a case of transfer of the passive formation rule from English to German; to what extent this is due to an overgeneralization of the German rule on the basis of the formal similarity of the languages is, again, difficult to determine.

The following items contained a verb, a 'false friend' created for this purpose that did not exist in the target language, at least not in the meaning it had in the source sentences. Yet, the learners seemed to be quite willing to be led down the 'garden path' to accept such translations. Thus, item

*412 They rushed to the station
Sie raschten zum Bahnhof

was accepted by 85% of the Finns very fast. In the Swedish-language Block 6,

*610 I make strong coffee
Jag makar starkt kaffe

was accepted by 35% of the Finns, while item

*609 Lisa rushed to the park
Lisa rasade till parken

was quickly accepted by 90% of the Finns.

Even item 403 containing the structure *verlassen für* + the expression of destination, which does not exist in German,

*403 I have left for Paris
Ich habe für Paris verlassen

was accepted relatively fast by as many as 32% of the Finns. There is no equivalent structure in Finnish, Swedish or German. German uses the verb *verlassen* 'to leave' with an accusative object to express, in most cases, the place which is left; both English (*leave*) and Swedish (*lämna*) use an identical structure (with a direct object) to express the same notion. However, it is typical of English to express the destination with the preposition *for*; among other verbs of movement which use the preposition *for* to express destination are *set out*, and figuratively used *head* and *make*.

Finnish normally translates the verb *leave* as *lähteä*, which requires the use of the elative case to express the place which is left (an equivalent of the preposition *from* on this occasion) and, similarly to German and Swedish, a case whose equivalent in English would be the preposition *to* to express destination.

The Finnish language learners seemed to have mastered the English structure well. The Finns' judgments were in line with the native speaker judgments: 85% Finns accepted the correct sentence 105 *He has left for Bolton*. The Finns' responses to sentence 106 **My friend left to Wales* were slightly more divided: four of them accepted the sentence. The results seem to indicate, though, that the Finns had learnt the construction well and may even have been aware of the markedness of this structure in English. The correct structure

seems to have caused difficulties for Germans in particular. Only 55% German university students of English and 61.5% of the German teachers accepted the correct sentence. On the other hand, 55% of the German students and seven German teachers considered sentence 106 correct.

Similarly, when the same structure appeared embedded in a Swedish sentence, the Finns revealed a tendency to accept the structure: item

*606 They have left for Dublin
De har lämnat för Dublin

was accepted by 40% of the Finns, just like sentence 711 **De har lämnat för Torneå.*[1]

It seems that the Finns have internalized the use of this English structure fairly well, better than the Germans. It may well be that the Germans, who speak a language which has a structure in some respects very similar to English but which differs from it as regards the essential syntactic relationships, are more easily misled by the formal similarity of the verb itself and tend to reject the form that they regard as unusual. Interestingly, the Finns displayed a great tendency to accept the English-based structure in the other two Germanic languages, which they did not know so well even though their L1, Finnish, uses a structure expressing destination similar to the one in German and Swedish. The question still remains whether this is an instance of lexical or grammatical transfer from the learner's L2 (English) into another non-native language. It may be more productive to regard this as lexical transfer. The existence of cognates, especially the formal similarity of the prepositions and the semantic similarity of the verbs in all these languages may have affected the Finnish learners' judgments and led them to transfer this structure, even though it can be regarded as linguistically "marked" in English, to other non-native languages perceived to be similar to English.

The adjective *interesting* has formal equivalents in German and Swedish. It also occurs very frequently in many types of discourse, and it is very likely that learners are familiar with it; in other words, it can be assumed that this structure is to a certain extent automatized in the language knowledge of advanced students of English.

All Finns (as well as all native speakers of English) recognised and accepted the correct sentence 125 *Are you interested in cars?* quite quickly; 95% of the German students accepted the sentence; moreover, they were almost as quick as the native speakers of English. In contrast, the erroneous sentence 125 **Is she interested of cooking?* was quickly and unanimously rejected by all native speakers of English; similarly, 95% of the Finnish and German students of English, along with 12 out of the 13 German teachers, found the sentence ungrammatical.

Item 406 below, which had the German preposition *in* for the correct *an*,

*406 Eve was interested in cars
Eve war interessiert in Autos

was accepted by over 50% of the Finns, but their slightly longer than avarage mean RT for this item indicates that the learners encountered problems in the process of judging the translation. In the Swedish-language block, the tendency to accept constructions with the

---

[1] The present writer was told by a native speaker of Swedish that he had actually heard this structure in certain parts of Sweden; it is regarded as an anglicism in Swedish, however.

preposition *in* was equally apparent: 60% of the Finns accepted the incorrect translation in item

*607 Len is interested in films
Len är intresserad i filmer

but, again, they spent more than the average time to judge this pair. In contrast, 11 out of the 13 German teachers rejected the German translation, while 65% of the German students of English accepted the translation. The result is perhaps easier to explain by reference to their relatively low mean RT, 2849 msec, for this item. Consistent with the overall results, this result shows that, again, the responses of the German students were faster but not nearly as accurate as those of the German teachers.

The same prepositional structure in sentence 510 **Er war interessiert in Katzen* was accepted by 60% of the Finns. Even more remarkable is the fact that 55% of them actually rejected the correct sentence 509 *Ich war interessiert an Tieren*. (Interestingly, as many as four (25%) German students of English also accepted the incorrect sentence; the German teachers were much more unanimous in rejecting the erroneous form, since only one subject out of the 13 accepted it.) Here a clear difference emerged between those Finnish university students who also studied German and those who did not, as 8 out of the 10 students of German accepted the correct form; in contrast, only one subject who did not study German found it correct. (Moreover, as many as eight non-students of German accepted the erroneous form). This result is in accordance with what has been noted before about transfer: language learners with a relatively low level of proficiency tend to have recourse to it more often than more advanced students of a language (see eg. Taylor 1975).

In Swedish, a similar tendency appeared: 50% of the Finns accepted the erroneous form in sentence 713 **Han är intresserad i fåglar*. Again, the level of language proficiency seems to play a part: of 15 subjects who did not study Swedish at the university, as many as nine accepted the incorrect sentence while only one student of Swedish did so.

The item containing the analytic comparative *mehr* + the adjective for the correct inflected comparative form *pünktlicher*

*414 She is more punctual than you
Sie ist mehr pünktlicher als du

was accepted by as many as 85% of the Finns. However, it obviously caused some problems for the learners as both the translation pair and the isolated German sentences containing the same structure received mean RTs that were among the longest in their task blocks. The corresponding sentence 526 **Elsa ist mehr tapfer als ich* was accepted by 75% of the Finns.

To sum up, among the English/German translation pairs, items 407, 412, and even 403 were judged relatively fast, while 406 and 414 received long RTs. On the whole, the isolated incorrect German L3 sentences that were based on English L2 received longer RTs than the incorrect ones that were not based on English, which would seem to indicate that there was a great deal of uncertainty among the subjects as to the correct grammatical form. The only exception to this was item **Er raschte zur Bank*, which contained a 'false friend' based on the English verb *rush*, which received quite fast judgment. Similarly, in the Swedish-language Block 7, sentence 716 **Hon rasade till Stockman* was accepted by as many as 85% of the Finns while sentence 713 **Han är interesserad i fåglar* was accepted by 50% of the Finns. However, sentence 719 **Lars makar starkt kaffe* was accepted only by 20% of the Finns.

Interestingly, all these cases except one seem to involve a fast acceptance of a particular lexical pattern or instance, rather than a rule (cf. Logan 1988). The only exception appears to be the extension of the English passive structure to German: 512 **Ich wurde ein Buch gegeben* was accepted by 55% of the Finns. At the same time, it received the second longest RT in the German block: obviously, standing on its own it caused some problems for the learners. Some sort of analogous process can have been at work, but it still seems to be the only clear case of grammatical transfer in the data. The rest of the cases seem to offer evidence of the indeterminate state of learner language as regards the structures in question, many for intralingual reasons (such as infinitive constructions).

8.7.3 Hypothesis 3: Learner responses to English L2 sentences

Above, attention was called to the way in which the effect of transfer might manifest itself during the task type under discussion. The reaction times were considered to reflect the degree of automatization of a particular linguistic structure. If the learner of a language quickly accepts a deviant structure containing an error based on L1 or any other language with sufficiently automatized access to relevant knowledge structures, it can be assumed that she has been assisted in the decision-making process by the existence of such ready-made subroutines. In Experiment 1, however, the widely different nature of the three learner groups made it difficult to compare the mean RTs for individual test items directly. For example, the overall difference in the processing speed of the Finnish and German university students make such comparisons rather uninformative in other respects. The other German subject group did not have a similar background in English language studies; their English language proficiency was quite varied and generally lower. Again, it was possible to compare items on the basis of the relative speed of the judgments and thereby establish a kind of processing difficulty hierarchy for the English test sentences across the subject groups. The sentences that were judged faster than the mean RT for the whole Block 1 were taken to be an indication of the use of faster, more automatized processes.

First, the correct sentences were generally judged faster than the deviant ones, although the native speakers did not display this tendency to the same degree as the language learners did. Still, even the NSs were slower to judge sentences that were deviant or otherwise unusual. This result supports the findings of Sajavaara & Lehtonen (1983); see also Sutter and Johnson (1990). The idiomatic use of articles, certain idioms, problems of word order, forms of strong verbs, were among the sentences that received the longest RTs. Very short RTs were obtained for sentences containing grammatical errors that were apparently so salient that they were immediately detected. The grammatical problems as such seemed to cover a wide range of linguistic phenomena.

The Finns were faster than average in accepting very few such erroneous items, though: among them were sentences such as 183 **They did fun of him* (accepted by 35% of the Finns), 140 **They pollute the nature* (accepted by 75% of the Finns), and 118 **We always travel with train* (accepted by 40% of the Finns). Of these, 183 could very well be the result of transfer from Finnish L1: the most usual translation equivalent of the verb *tehdä* in the Finnish idiom *tehdä pilaa jostakin* 'make fun of somebody' is *do*; sentence 140, and to a certain extent 118, however, contained structures that are subject to interlanguage system-internal variation. The German university students of English, on the other hand, showed a tendency quickly to accept items such as 152 **My father have to brothers* (accepted by 25%), 120 **Dan lives at the country* (accepted by 35%), 140 **They pollute*

the nature (accepted by 80%), 179 *(*)Larry cooks good coffee* (accepted by 70%), 118 **We always travel with train* (accepted by 30%), and 151 **Ted drink tea every morning* (accepted by 50%). The German students' responses to 118, 120, 179, 188 and perhaps to 140, as well, were influenced by their German L1, which has equivalent constructions. Their responses to 151 and 152 could be better explained as having a system-internal source. Figures 14 and 15 below show the learner and NS performance for sentences that have an equivalent structure in either Finnish L1 (183) or both German and Finnish L1s (118, 120 and 179).

118 *We always travel with train
120 *Dan lives at the country
179 (*)Larry cooks good coffee
183 *They did fun of him

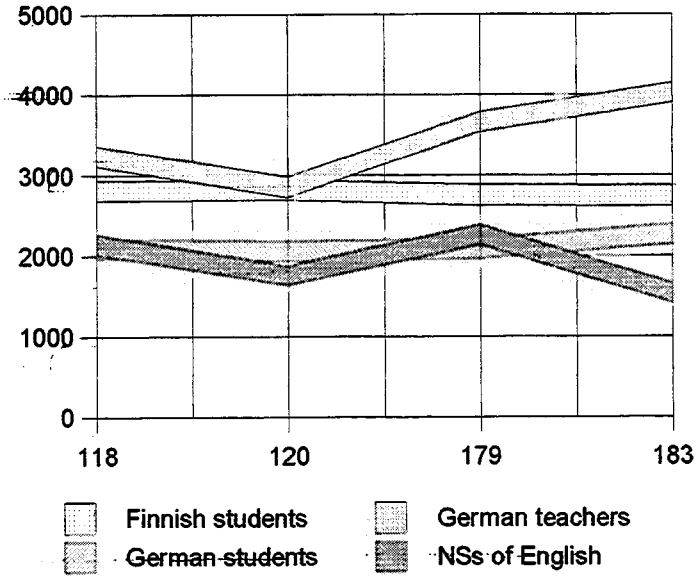


FIGURE 14 Learner and NSs judgment times for sentences receiving a fast judgment by Finnish and German learners of English.

118 *We always travel with train
120 *Dan lives at the country
179 (*)Larry cooks good coffee
183 *They did fun of him

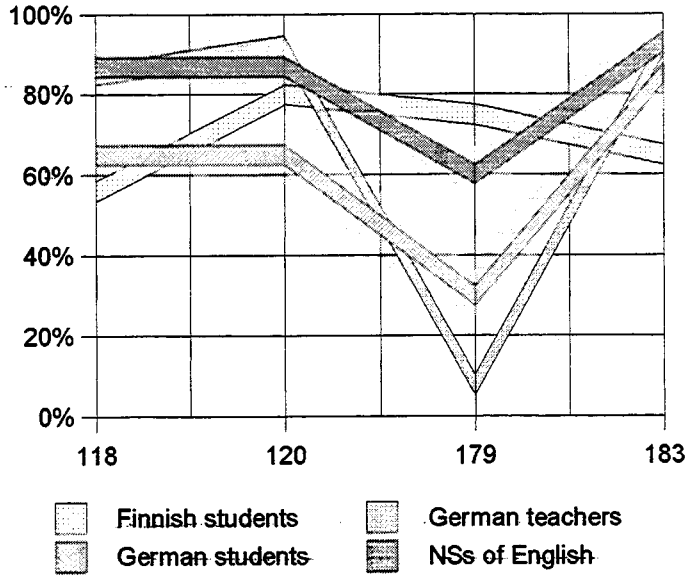


FIGURE 15 Learner and NSs judgments for sentences receiving fast judgment by German (sentences 118, 120 and 179) and Finnish (sentence 183) students of English.

The Finns were quite slow in judging structures involving, for example, the use of infinitives (164, 168) or involving the definite article (142). They were also quite slow in judging certain prepositional constructions (102, 106, 108, 116, 124, and even 126). This is quite significant since these structures were by and large modeled by their Finnish L1; in other words, the Finnish learners of English L2 were quite familiar with such structures in their interlanguage, unlike the German learners of English. Figures 16 and 17 below show the results for these six sentences.This is in accordance with what has been said above on the overall difficulty of such structures for Finns learning English. Furthermore, in all such cases, except 164, 168 and 142, which involved the use of grammatical categories that had no equivalent in Finnish, the sentences seemed to have a translation equivalent of the deviant structure in their L1 counterparts. This seems to suggest that, first, the meaning of the

102 *The price depends of it
106 *My friend left to Wales
108 *Jim found it from the shop
116 *He left it to Jill's room
124 *He lived on the 12th century
126 *Is she interested of cooking?

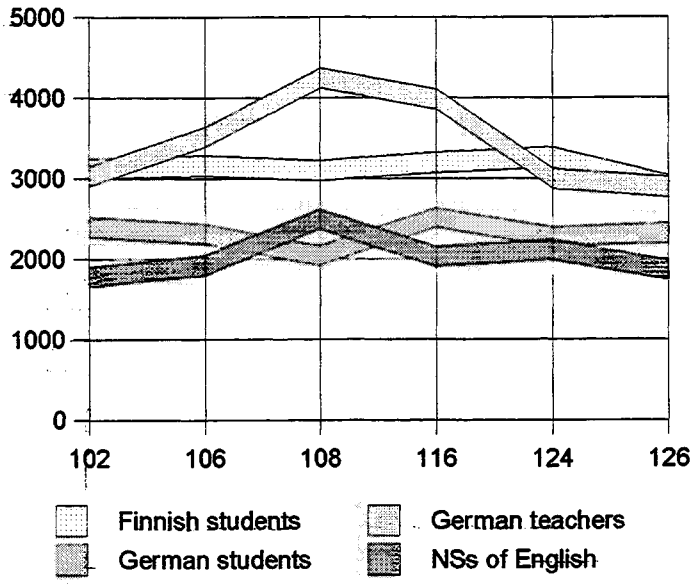


FIGURE 16 Learner and NS judgment times on six sentences modeled on Finnish.

102 *The price depends of it
106 *My friend left to Wales
108 *Jim found it from the shop
116 *He left it to Jill's room
124 *He lived on the 12th century
126 *Is she interested of cooking?

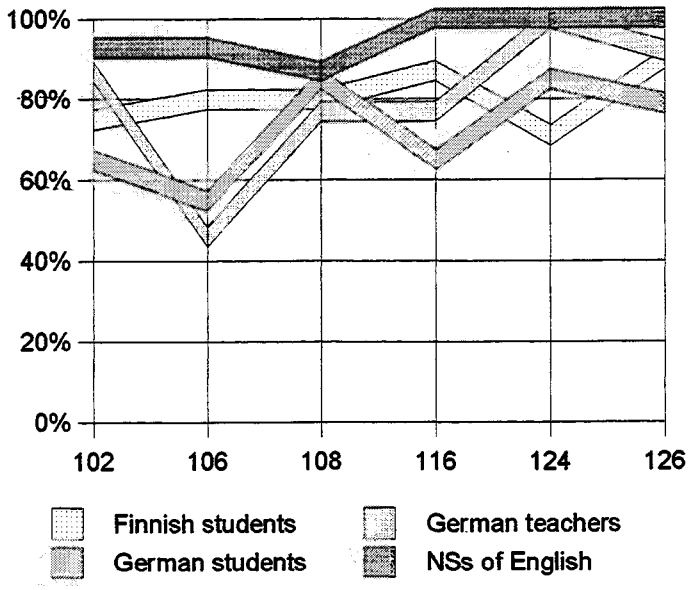


FIGURE 17 Learner and NS judgments on six entences modeled on Finnish.

sentence together with the presence of the preposition activated L1 knowledge. This very likely caused them to start monitoring, presumably because they were aware of the difference between Finnish and English in these cases. This seems to indicate something about the way such knowledge has been acquired, organized, and now accessed during second language performance. Second, the Finnish learners tended to have to search their interlanguage knowledge for such items as the *to* infinitive and the use of the definite article. The *to* + infinitive, the articles and the 3rd person *s* are a common problem for many

140 *They pollute the nature
142 *He lived in the hot Morocco
151 *Ted drink tea every morning
152 *My father have two brothers
164 *He heard Jeff to come in
166 *We only try help them
168 *George made her to steal it

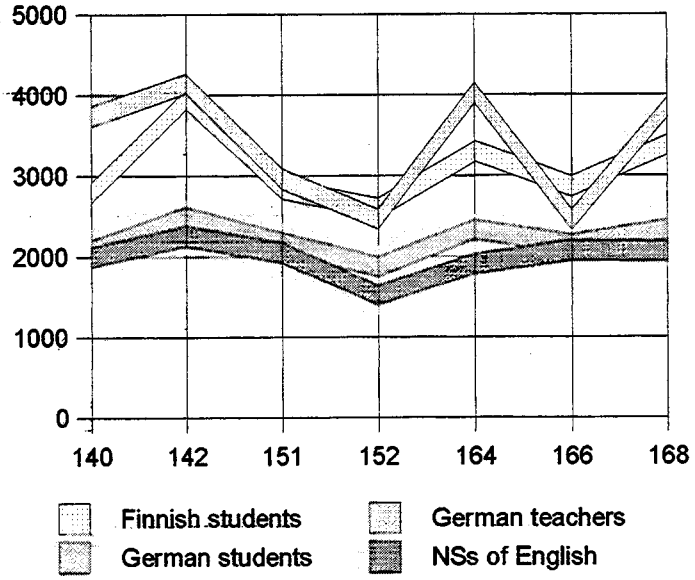


FIGURE 18 Learner and NS judgment times for sentences with system-internal errors.

140 *They pollute the nature
142 *He lived in the hot Morocco
151 *Ted drink tea every morning
152 *My father have two brothers
164 *He heard Jeff to come in
166 *We only try help them
168 *George made her to steal it

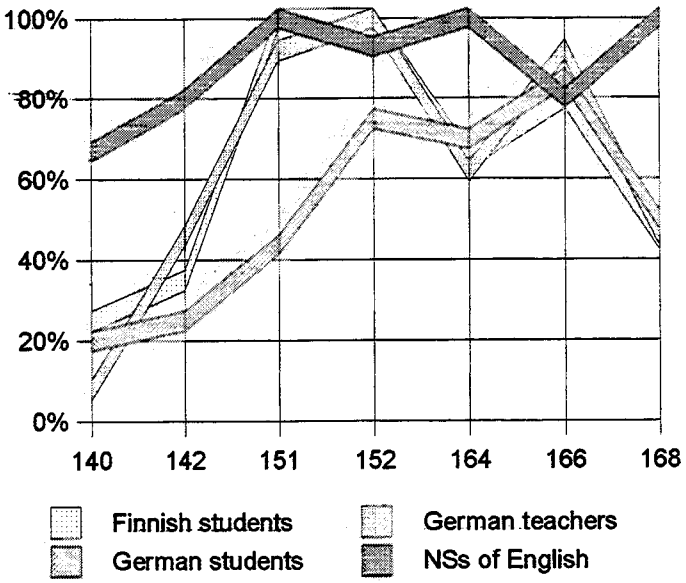


FIGURE 19 Learner and NS judgments on sentences containing system-internal errors.

learners of English (see eg. Johnson and Newport 1989). They are phenomena that are subject to system-internal variation in English. There are no direct equivalents in Finnish for these grammatical structures; in languages typologically similar to English, the L1-based system could be of help. However, the results showed a great deal of variation among all learner groups as to the use of these structures. Figures 18 and 19 show learner ans NS performance for these sentences.

A common feature of English, German and Swedish is the use of the infinitive as a complement of the verb with a special marker, which has no equivalent in Finnish. English verbs differ in the form of the infinitive (with or without *to*) they allow. Not all the items included in the experiment were particularly successful. There was a set of sentences, however, where certain differences in the performance between the learner groups seemed to emerge.

For example, consider the following pair of English sentences: The NS judgments on 168 **George made her to steal it* and 167 *I made him drink it* were practically undivided. All Finns accepted 167 as well. The responses of the Germans showed a great deal of variation, however: 46% of the German teachers and 40% of the students rejected the grammatical sentence. All NSs of English accepted the grammatical sentence 165 *He never tries to stop me.* The Finns were also quite unanimous in their acceptance of the sentence. Sentence 166 **We only try help them*, however, was accepted by three NSs, which seems to indicate that there is some variation in English as to the acceptable form of the infinitive with the verb *try*. There were a few learners who also accepted the sentence.

The learner judgments on the infinitive forms of English varied to a greater extent among the Germans. Interestingly, the Finns were all able to recognize the correct use of the infinitive accurately and very quickly, as the sentences received quite short RTs. The Finns' and Germans' prolonged RTs in the case of 168 suggest that they had to resort to more controlled processing and try to actively search for relevant information to make their decision, which may or may not have been correct. Moreover, the Germans seemed to be unsure of the correct form in 167, and to a certain extent also in 165. In the case of 167, the result may have been caused by the combined effect of various factors: first, there is no formal equivalent in German of the verb *make* in the sense used here, which means that the German learners would find no previous knowledge in their L1 to which 'anchor' this new usage. Instead, they would have to try to mark the relevant syntactic information in their interlanguage lexical entry *make* and eventually build a sufficiently automatized subroutine for a quick processing of the linguistic item, which takes time. In a task that requires quick decisions and which therefore limits the time available for processing, the learner may in fact resort to more automatized routines, ie. the acceptance of the particle *to* in verb complements and the rejection of cases where it is absent. Second, the existence of the formal equivalent in their L1 (the particle *zu*) with rules of its own may have caused some further confusion, either during the performance or during the learning process itself: it is not inconceivable that some of the subroutines connected with the use of the infinitives in their L1 become activated during the processing of a language sufficiently similar.

The 3rd person singular *-s* is problematic to many learners of English. As Dagut and Laufer (1982: 31) point out, the source of this error, sometimes referred to as "the 3rd person singular *-s* problem", has presented difficulty for linguists. The errors made by language learners concerning the use of the third person singular present tense *-s* has often been pointed out as evidence against L1 transfer. It has been argued that if the learners' L1 had an effect on their developing interlanguage system, this would be evident in the different nature of errors made as regards the use of *-s*. Since this does not seem to be the case, it

is usually explained as overgeneralisation of the dominantly non-inflected English present tense form. Dagut and Laufer, however, claim that it can be explained by what they call indirect interference, ie. interference which occurs when the learner tries to deal with an L2 category non-existent (or virtually non-existent) in his L1. The learner is then forced to resort to false analogies, or on mere guessing. Interestingly, they point out that between English and Hebrew, such interference would arise because of the completely different inflexional systems in the two languages: as Hebrew marks the verbs for gender, person and number, it is difficult for the Hebrew language learner to understand why she should now distinguish only the person (and only 3rd person singular). Similarly, of course, it could be argued that it does not make sense for a Finnish (or German) language learner to distinguish only between the 3rd person and the other persons. On the other hand, Johnson and Newport (1989) report on their study of the occurrence of these phenomena in the English L2 of Chinese L1 immigrants in the United States.

In Experiment 1, there were two sentences containg an error in number concord. Both Finnish and German (but not Swedish) are more distinctive in their use of inflectional endings. Consequently, one would expect the learners with such an L1 to be more aware of them in the target language. However, over-generalization seems to play a part instead; the omission of the singular *-s* is not an infrequent error among the Finnish learners of English, at least among the beginners. Of the four subject groups, the German students were the least accurate in their judgments: whereas all native speakers, the Finns, and all German teachers except one rejected the sentence 151 **Ted drink tea every morning*, 50% of the German students actually accepted the sentence. The same tendency is apparent in sentence 152 **My father have two brothers*: whereas all the other subjects (except one native speaker) rejected this sentence, 25% of the German students found nothing wrong with it.

One explanation for the difference between the performance of the Finnish and German students of English could be that the Finns were more aware of the differences between their L1 and English. This raised level of awareness might have been in effect during the judgment task, but it could also have affected the acquisition process and the representation of knowledge thus gained. Jordens (1977), for example, speaks of 'suspicion-inducing influence of teaching' (Kellerman 1977). The encountering of such pieces of knowledge, warning 'flags', may increase the capacity needed for processing as more resources become activated (cf. Logan 1985: 376-377), which, again, would reduce the speed of processing. On the other hand, the learners might have become so skilled in dealing with often-encountered linguistic structures that they would be able to process them quickly. In the cases described above, both Finnish and German learners show fast reaction times when dealing with linguistic material that contained errors that they might be reasonably familiar with in their interlanguage, and been warned against not to use, either by themselves or the teacher.

In reference to what Logan (1985) noted on the difference between skill and automaticity, it can be that the Finns' language skills were high enough for them to be able to control their performance to the extent that it appeared nevertheless quite automatic for the sentences that they were able to judge quickly, hence the short reaction times. However, the Finns' longer RTs in general indicate that they were more inclined to monitor their performance, similarly to the German teachers. Both of these groups scored better than the German students did. Therefore, it appears that learner characteristics very likely influence the occurence or non-occurrence of transfer.

However, the Finns' longer reaction times seem very likely to have been also a function of the typological dissimilarity between Finnish and English. Such results as those presented above support the conclusion that the German students were willing to utilize the more automatized processes based on their L1 in their analysis of English sentences. The Finns, on the other hand, even if their L1 supported the relevant structure in English, were inhibited in their application of L1 knowledge. This appears to be not only an effect of the greater perceived distance between Finnish and English, but also of the real distance between the languages. The kind of instruction learners had received may also have played a role: the Finns may have been more 'suspicious' of L1-based structures in L2. In addition, it could even be that the way the Finns approached English may reflect underlying differences in how for example graphemes are utilized in their L1 (cf. Lehtonen 1983). Even if the Germans felt that the connection between their L1 and English was quite remote, the underlying typological similarities of which they might not even have been aware (it is, after all, far more likely that they had paid attention only to existing differences between the languages in language learning situation) influenced their performance.

8.7.4 Individual differences in learner performance

Interesting data were provided by the individual language learners' performance in various task groups. This became especially clear in the case of the Finns who carried out tasks in four different languages. Table 5 below shows the correlation coefficients between the mean RTs of the Finns for each of the language blocks. As a matter of fact, in each of the

TABLE 5 The correlation coefficients between the Finns' mean RTs in the various language blocks and their statistical significance. ** = $p \leq .01$.

	Finnish	English	German	Swedish	English/ Finnish	English/ German	English/ Swedish
Finnish	-	.82**	.59**	.75**	.73**	.52**	.78**
English		-	.72**	.75**	.71**	. 74**	.85**
German			-	.78**	.60**	.91**	.82**
Swedish				-	.71**	.68**	.89**
English / Finnish					-	.58**	.74**
English / German						-	.74**
English / Swedish							-

three learner groups there seemed to emerge a set of subjects who were consistently slow in their performance across the languages. These language learners were apparently controlling their performance to a greater extent than the other subjects in the group. Correspondingly, there also emerged language learners who were quite consistently fast in their responses, as shown in Figure 20. There appeared to be more consistency at the 'slow' end of the scale, as far as the performance in the various blocks is concerned. The subjects' overall speed seemed to depend on factors such as familiarity with the language: for example students of German tended to be faster in German than the non-students.

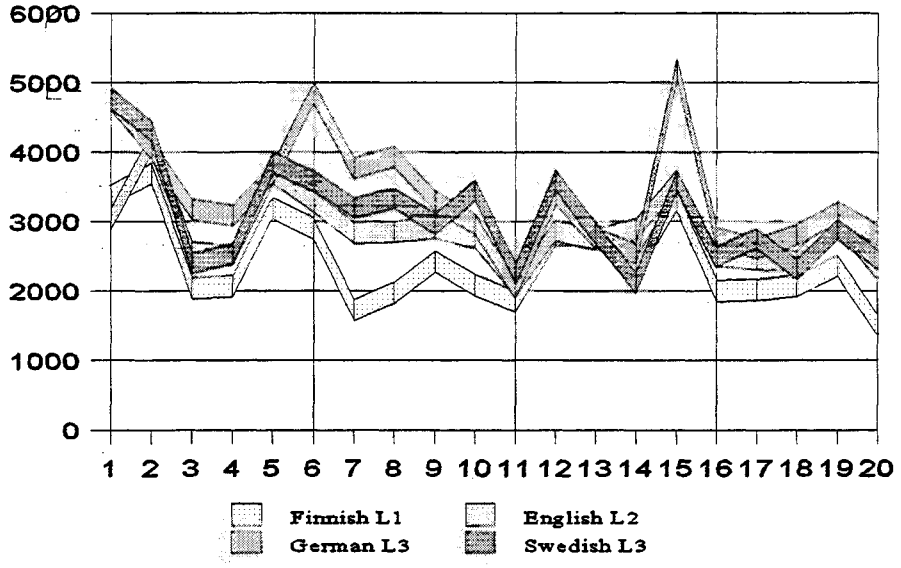


FIGURE 20 Mean RTs for individual Finnish learners of English L2.

The results show that those who were slower in making decisions in their L1 were also inclined to be slow in a foreign language (although they were exceptions). This may be due to a tendency to monitor or simply an indication of a difference in 'processing styles' (cf. Lehtonen and Sajavaara 1985). The processing style of a subject seemed to affect his or her willingness to transfer as well: a 'fast' language learner (eg. subject 11) accepted English-based expressions in German and Swedish very fast. Thus she was quite able to switch over to more automatized routines whether they had their source in L1 or L2 (as in this instance seemed to be the case). However, in support of what was said about the 'markedness' of Finnish, it appeared that Subject 11 was not at all so willing to accept structures that seemed to be based in Finnish and was usually even able to reject them very quickly. Such between-subject variation in processing speed may be a reflection of individual differences that have their basis in cognition and even personality (cf. Lehtonen and Sajavaara 1985).

As the processing of one's native language is usually regarded as highly automatized, it is not likely that a person who is controlling her performance in her L1 would use automatized processing in L2. On the other hand, the number of the L1 sentences included in the experiment is not sufficient for far-reaching conclusions. It may also have been that L1 material of this study, being rather ambiguous in meaning, could have caused a similar activation of linguistic knowledge in the Finnish subjects as foreign languages seemed to trigger. This could also explain why it was the English block that correlated most consistently with all the other language blocks: the Finns were reasonably proficient in English but not to the extent that they would have totally relied on their automatic routines; hence the use of the controlled processing mode. It also indicates an involvement of metalinguistic skill that is somehow tied to second/foreign language learning or processing, an experience which usually results in a greater degree of awareness of and attention to linguistic form.

Havia (1982: 106) noted that an individual's RT was not a good indicator of her language proficiency. However, since he used the subjects' university grades for certain English courses as the measure of language proficiency, his findings could have been simply

a result of inadequate instrumentation. (On the other hand, there is no apparent connection between the learners' speed and accuracy since the correlation coefficients were all statistically insignificant.) However, it is obvious that the level of language proficiency must play a part in the performance of these tasks: not only were the NSs significantly faster and more accurate than the learners in judging sentences, but there was also a difference among the Finns that could only be explained by language proficiency.

8.7.5 The effect of language proficiency

The effect of language proficiency in this experiment is three-fold: First, it can be assumed to underlie the difference in performance between the NSs and the learners of English and German, respectively. As became apparent in Chapter 7.7.1 above, the NSs of English were significantly faster and/or more accurate than the Finnish and German learners of English. Similarly, the NSs of German were significantly faster in their judgments of German than the Finns were. Secondly, the Finnish students' varying levels of proficiency in German must explain some of the differences found in their performance in German L3. Figure 21 below shows the mean RTs for both of these groups in all language blocks. The subjects' German proficiency was not formally tested; instead, the Finnish subjects were divided into two *post hoc* groups: those who studied German as a minor, and those who did not.

The average judgment times of the students and non-students of German L3 in the language groups that do not involve German are practically identical. There is a statistically significant difference between their performance in the two German-language bocks, the

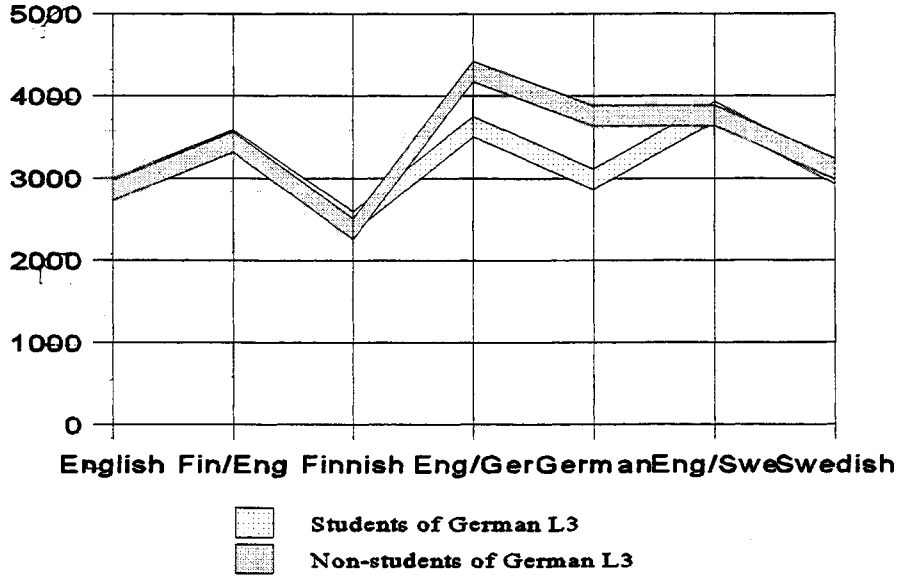


FIGURE 21 The mean RTs of students (n=10) and non-students (n=10) of German in different language blocks. subject at the university

students of German being clearly faster in their judgments. These results can only be explained by their greater familiarity with the language. In other words, they had greater procedural skills in dealing with German-language input than their peers who did not study German.

Thirdly, the German teachers, who had not studied English at the university (apart from one subject), were significantly slower than the German students of English. However, no such difference emerged in the accuracy of their judgments. As mentioned above, the German students were remarkably inaccurate in their judgments of English sentences.

8.8. Conclusion

The aim of the study was to attempt to find evidence for influence between two non-native languages in the acceptability judgments made by Finnish university students of English on German and Swedish sentences. As it turned out, the influence was the strongest in the case of lexical items such as false friends. Certain prepositions also seemed to be relatively transferable; whether the acceptances were caused by overgeneralisations within the interlanguage system or by transfer from L2 is difficult to determine. It is likely, however, that the formal similarity between English and German (or Swedish) items may have been enough to activate the subjects' L2 knowledge. According to Ringbom (1982: 94), the organization of the lexicon on the basis of formal similarity instead of semantic network associations may be a characteristic of beginning learners. Since the Finns were very advanced learners of English, they may have organized their English L2 lexicon for the most part on the basis of the meaning of lexical items; their German and Swedish lexicons may have been organized on the basis of formal similarity, which may have also caused close associations with the English lexical entries. Moreover, most cases of cross-linguistic influence between non-native languages seem to have been caused by formal similarities between the source language and target language (cf. Ringbom 1986).

What seemed to be a case of rather rare grammatical transfer between non-native languages concerned the carrying over of the rules of English passive formation into German. The Finns' decisions could have been influenced by training and the frequency of this particular passive structure in the L2 input: English L2 passive sentences where the object of the corresponding active sentence has been moved to the subject position (instead of the indirect object) are much more common, but in the case of the verbs with both the direct and indirect object it is more usual for the indirect object to function as the subject of the passive sentence. When there was no equivalent structure in the learner's L1, such as articles and infinitive markers, the Finns' decisions seemed almost random, especially in the languages with which they were less familiar. The influence of English seemed rather weak in these areas; on the contrary, concerning the use of the definite article the English usage seemed to be influenced by the other non-native languages German and Swedish, which both use the definite article differently.

It also became apparent that the three learner groups each processed the English sentences differently, and that this seemed to be somehow bound with the subjects' L1. The German students processed the sentences very fast, almost as fast as the native speakers. In contrast, the German teachers and the Finns were much slower; for some sentences even, the Finns' and the German teachers' performance profile was quite similar to each other. It appeared that the Finns, however, were far more accurate in their judgments than the German students. This seemed to suggest that the Finns tended to monitor their performance in processing English more than the German university students of English did, and that the German teachers controlled their performance to a much greater degree as well. The longer

mean RTs for the Finns may also reflect the differential nature of language processing strategies.

There was also evidence for the operation of some notion of 'markedness' whether psycholinguistically motivated (cf. Kellerman 1977) or based on typological universals (cf. Eckman 1977). The Finns' responses may have been affected by the greater perceived distance between Finnish and English, since they refused to transfer L1 structures even though they may have been less 'marked', universal, or otherwise language-neutral. For example, most of the Finns rejected sentence 106 **My friend left to Wales,* whereas the German students accepted it and rejected the correct sentence.

Ringbom (1982,1987) has shown how the Swedish-speaking Finns have a better start than the Finnish-speaking Finns in learning English because of the typological similarity between their L1 and L2. The results of the present experiment suggest that the Finns have a clear disadvantage at least in the processing of English as compared to the Germans who took part in this experiment. However, the greater accuracy of the Finns seems to indicate that their English language proficiency is better. Moreover, the relative automaticity in the processing style of certain Finns implies that it is indeed possible for a Finn to achieve a very high degree of automaticity in English. It may even be that because they cannot rely on their L1 in language learning they have to start the building of their interlanguage with a lesser degree of transfer from L1 knowledge or language processes. Accordingly, they may avoid later fossilizations that could be due to direct transfer and application of automatized routines from L1 to L2. Such automatized procedures are not easily changed (cf. Shiffrin and Schneider 1977). This interpretation finds support for example in Ringbom (1987), who notes that the English language proficiency of the Finns may actually surpass that of the Swedes when high enough a level is reached.

There also emerged some fairly consistent individual differences in the processing of the sentences. Certain subjects seemed to be consistently slow in performing such tasks as these, regardless of the language involved. This raises a question whether the differences might be somehow related to cognitive and affective factors. The low number and ambiguous nature of some of the L1 sentences in the present experiment affects the reliability and validity of the results, however. The following experiment was designed to investigate, first, whether the overall individual differences found in Experiment 1 would appear with another set of L1 and L2 sentences, and second, whether the individual differences could be related to certain cognitive and affective constructs.

9 Experiment 2

9.1 Research questions

In Experiment 2, an attempt was made to account for individual variation across reaction time latencies in acceptability judgments by Finnish advanced learners of English. As in the previous experiment, the subjects were asked to judge English sentences that contained errors caused by the influence of their native language. The purpose of this experiment was to try to relate the acceptability judgments and judgment times by Finnish advanced learners of English to a number of cognitive variables Again, within the framework of the present experiment, reaction times are assumed to reflect the degree of automaticity in the way the knowledge is accessed (Sajavaara and Lehtonen 1983,1989; Lehtonen 1989).

It was assumed that the type of task involved would comprise at least the following aspects of second language performance that should somehow be accounted for: automatic and controlled processes and their link to the role of declarative and procedural knowledge (including the language learner's implicit L1 knowledge, implicit interlanguage knowledge, and explicit knowledge of L1 and L2), potential transfer of L1 procedures and structures and the language learner's metalinguistic awareness/skills. The starting hypothesis was that acceptability judgments would reflect the influence of L1 in that Finnish language learners would show a tendency to accept structures in English L2 containing an error based on L1 structures. To compare the Finns' performance with that of native speakers', acceptability judgments were obtained from five native speakers of English. On the basis of earlier studies (see eg. Lehtonen and Sajavaara 1983, McLeod and McLaughlin 1986: 112; see also the previous experiment), it was assumed that non-native speakers require more time in making such judgments.

The experiment was primarily focused on discovering patterns in learner performance. It was hypothesized that various cognitive variables such as field-dependence-independence would correlate with the degree of automatized processing as reflected in reaction time latencies and possibly also the accuracy of the judgments themselves. The other variables included in the study were the extent of L2 communicative anxiety felt by the language learners, verbal reasoning in the native language, and their verbal and non-verbal thinking styles.

The cognitive and affective variables were chosen on the basis of studies investigating the "good language learner" (se eg. Naiman et al. 1978), metalinguistic awareness (see eg. Masny and d'Anglejean 1985), various cognitive styles (Witkin et al. 1971), and communicative anxiety (see eg. Lehtonen, Sajavaara and Manninen 1985). The tentative and hypothetical nature of many of these psychological constructs and their operationalizations was recognized even though they had been used before in linguistic or psychological research.

9.2 Subjects

Two groups of subjects were used in the experiment. The Finnish language learner group consisted of 21 students of English. Originally, the Finnish group consisted of 23 subjects, but two of them, Ss 9 and 23, were excluded from the statistical analysis of the data. They are, however, included in the study of the individual subjects as they offer an interesting point of comparison: S 9 could be characterized as fully bilingual in Finnish and English since she had lived for seven years in an English-speaking country and was partly educated there; S 23 was a Swedish-speaking Finn: she reported Finnish to be her stronger language but her L1 was Swedish. For these reasons, it was considered wiser to eliminate these two bilinguals from the overall analysis of the Finns' performance which was to be compared with the performance of the NSs. Eight of them (Ss 1,2,4,14,16,18,19, and 20) were first-year students of English at the English Department of the University of Jyväskylä, eight were second-year students (one of them had French as her main subject) (Ss 3,6,7,11,12,17,21; S 5 was a fourth- year student of French), and four were third-year students (Ss 8,10,13,and 15). All Finnish subjects took part in all the tests in the study and also filled in both the background and thinking style questionnaires. The group of native speakers of English (NSs) consisted of 5 university teachers of English. Two of them had only worked for approximately three months in Finland; in the case of the NSs, more important than the distinction between nationalities was the difference in teaching experience, in particular the familiarity with errors made by Finnish learners of English. The individual differences between the NSs in reaction time latencies can at least partly be explained by this fact. The NSs of English took part in the grammaticality judgement test only.

Two of the Finns (Ss 13 and 16) were actively studying Swedish at the university. Six Finns (Ss 3,4,7,9,15,and 18) had stayed more than six months in an English speaking country. All subjects knew at least one foreign language in addition to English and Swedish.

No attempt was made to control the subjects' proficiency in English, nor was their level of proficiency in Swedish controlled in any way; it was assumed that all subjects as students of English at the university level (and as having passed the entrance examination) had a sufficient knowledge of English to be qualified as advanced learners of English. Swedish, being the second official language in Finland, is a compulsory subject in the Finnish school system, so the subjects were expected to have some knowledge of it.

9.3. Variables

The experiment involved two types of dependent variables, reaction time to and grammatical acceptability judgments on isolated sentences in three languages. The kind of reaction time obtained from such a decision task is global (see eg. Zwitserlood 1989): the task comprises several processes. It was assumed that a fast rate of decision-making reflects a higher degree of availability of knowledge and degree of automaticity in the accessing of this knowledge (see eg. Sajavaara and Lehtonen 1989: 42). In other words, reaction times were thought to reflect the extent to which the subjects were able to use automatized processes in judging these sentences.

The study included several cognitive variables: The first variable, the field-dependence-independence dimension of cognitive style, is a well-known concept in differential psychology. This notion has a long history of research (Witkin et al. 1971, Witkin et al. 1977a), and it has well-established methods of research. It was tested by using a standardized Embedded Figures Test (EFT).

The second group of variables chosen was closely connected with communicative apprehension in second language communicative situations. There has been a certain amount of research in this area in the field of native language communication, but its application to second or foreign language research has been less common (see eg. MacIntyre and Gardner 1991). A questionnaire was used to measure this dimension; the underlying assumption was that it is possible to operationalize this rather elusive construct in this manner. The items in the questionnaire were based on a communicative anxiety questionnaire (CAQ), where an attempt was made to measure Finnish language learners' communicative anxiety in English as a second language (Lehtonen, Sajavaara and Manninen 1985). Only the items with high loadings in such factors as social anxiety, anxiety about errors, reliance on English proficiency, and even a few items on the degree of reward given in English classes at school were included in the questionnaire. The format used in the questionnaire was an end-defined Likert-type scale. The items formed six groups that gave the following variables:

REWENG:	the amount of reward given in an English class
ANXERR:	anxiety about linguistic errors
SOCANX:	social anxiety
RELPROF:	reliance on language proficiency
ENGPROF:	self-evaluation of proficiency in English in both modes of production and comprehension; this was merged with RELPROF in the statistical analysis of the data.
SWEPROF:	self-evaluation of proficiency in Swedish in both modes of production and comprehension).

The value for each variable was the mean score of the answers for each factor. It was assumed that the learners scoring low on social anxiety or error anxiety scales of the questionnaire, ie. subjects who felt themselves to be more apprehensive in communicative situations, would perform more slowly. The underlying assumption was that such scores functioned as indicators of a general attitude that would give rise to the increased use of control in the test situation.

The third variable chosen involved the notion of verbal and non-verbal thinking styles. This notion is closely connected with the work of Paivio. The Individual Differences Questionnaire (IDQ) (published by Paivio and Harsman 1983) developed by him and his collaborators is used here without any commitment to the theoretical framework of the role of imagery in cognition and language processing (the Dual Code Hypothesis). The reason for using this is that it was seen as a relatively widely used method of measuring individual verbal and imaginal habits. It also seems to be quite effective in differentiating between subjects. It was hypothesized that at least the factors to do with verbal ability would be related to reaction time latencies. The questionnaire used in the present study consisted of 93 items designed to measure individual differences in verbal and imaginal habits and thinking styles. The original questionnaire was used with certain modifications, however: first, seven items were added from the questionnaire used by Spitzer (1986); secondly, the format of the questionnaire was changed, partly on the basis of Paivio and Harshman's recommendation (1983; see also Hiscock 1979) from True/False statements to the end-defined Likert type scale format. (An end-defined scale is a Likert scale with only the ends

labeled 'strongly agree' and 'strongly disagree'; there is another type with clearly defined alternatives such as 1-2-3-4-5 from 'strongly agree' to 'strongly disagree'.) All items were translated into Finnish. On the basis of the questionnaire, nominal scores of the following kinds were counted:

VERBAL: the nominal scores on the original verbal scale
IMAGIN: the nominal scores on the original imaginal scale

In addition, the items were also divided into seven factors gained from the factor analysis of the original questionnaire by Paivio and Harsman (1983) and Spitzer (1986).

PF1: 'good verbal ability'
PF2: 'habitual use of imagery'
PF3: 'concern with the nondeviant use of words'
PF4: 'reading difficulties', or rather absence of them as this variable was counted as verbal in content.
PF5: 'use of images to solve problems'
PF6: 'vividness of dreams, day dreams,and imagination'
PF7: 'verbal thinking'.

Although the original questionnaire was used to gather data, for reasons of greater reliability only the items that received high enough factor loadings (r>.30) in the original study were used in statistical analysis. The answers for these items falling in the area of 3, roughly equivalent of 'I can't say', were recorded as missing data and given the value of zero.

To measure the subjects' ability in verbal reasoning in L1, a standard test used by vocational guidance councelors was used. It resembles the cloze test in that it consists of a text with gaps (although not at regular intervals). The scores were standardized although both raw and standardized scores were used in calculating correlations (it turned out to have no bearing on the results). This variable was called V.

The subjects also filled in a short questionnaire on the background information that was thought to be of relevance, such as the length of studies in English or any other foreign language and stay in an English-speaking (or any other language) country.

9.4. Design of test sentences

The materials for the reaction time experiment consisted of eight pairs of ungrammatical sentences with their grammatically acceptable counterparts in English, Swedish, and Finnish. The particular grammatical problem chosen for the study concerned the use of certain prepositional constructions in English. These constructions were chosen on the basis of the potential difficulty they might present for Finnish learners of English. All these constructions have an equivalent surface structure in Finnish with the case ending *-ssa* or *-ssä* (the inessive case). For six of these, English uses the prepositions *on* or *at*; there were two constructions, however, containing expressions for which English uses the preposition *in*.

Within the various linguistic subsystems where transfer can be said to take place, prepositions create a specific problem; linguists have placed them now within syntax, now within lexis. Furthermore, as Stedje (1977: 150) points out, there are areas in which grammatical interference influences the level of lexis (or at least function words) in another language. For example, Finnish often expresses by means of grammatical elements what English and German realize by using lexical elements (or, if one prefers, function words),

that is, Finnish uses case endings to express what the other languages express with prepositions.

Prepositional constructions are therefore one instance of a grammatical phenomenon that has no similar formal equivalent in Finnish. Where other European languages use prepositions or postpositions, Finnish has case endings (see eg. Karlsson 1983: 13). As a matter of fact, Finnish seems to have developed in quite another direction than Indo-European languages, where case endings have given away to the expanding use of prepositions (Hakulinen 1979: 98). However, both case endings and prepositions have various grammatical and semantic functions of a similar kind to fulfill. On the one hand, this lack of formal correspondence between the learners' L1 and English L2 and Swedish L3, which again are typologically close to one another, provides a fruitful starting point for the study of transfer; any influence of the learner's mother tongue would in this case have to work through learning hypotheses and assumptions of translation equivalency between L1 case endings and various target language prepositions. The early ties formed between L1 and L2 would then affect the formation of the language learners' L2 interlanguage system. Naturally enough, even those languages that do have preposition systems of similar kind can never be quite identical in this respect. For instance, the use of prepositions in English and Swedish can differ considerably.

Thus, the most important question to be asked in connection of prepositions concerns the manner and stage at which transfer can take place when and if prepositional expressions are tranferred from language to language. The verification of transfer in the case of prepositional expressions is exceptionally difficult because of the interplay of various factors involved in their use. Errors seem to be the only way that transfer can be easily identified. However, such errors especially in the use of prepositions can also be attributed to various other sources, the most obvious of them being the effect of the interlanguage system itself (ie. intralingual errors). When there exists a clear distinction in function and/or form between two candidates, it is easier to say that the choice of the one rather than the other is at least partially influenced by another language system. For example, if a Finn chooses to say, **Rose had a ring in her finger* while meaning 'Rose had a ring on her finger', the product could be accounted for as a result of the potential influence of the speaker's L1: Finnish uses a case ending generally translated as *in* to express this type of relationship.

Another important question concerns the location of the linguistic information contained in the prepositions: is it part of a lexical entry's lemma information, or is it encoded as purely grammatical, procedural knowledge? It is probably more likely that many lexical items such as verbs and adjectives carry a list of prepositions as part of their lemma information to signal what preposition expresses what grammatical relation when used together with the lexical item, but this is by no means certain. Consequently, any transfer in the use of prepositions could be interpreted as a complex process where semantic-syntactic content of the L1 form is transferred either wholly or partially into the emerging interlanguage expression. Matters are even further complicated by the differences in the structure of mental lexicons between typologically such different languages as Finnish and English when one language uses case endings and the other prepositions. The inflectional forms seem to be listed in a lexical entry's lemma information as well (although there is no certainty about that either). How does the language learner arrange her interlanguage lexicon? No definitive answers to such questions have been found as yet. Such questions as these have to be borne in mind, however, when the data are dealt with even at the level of observed language behavior .

In the following discussion, reference will be made to the distinction made by Lyons (1969: 303) between local and formal use of prepositions. Local use includes place and time expressions whereas formal usage involves grammatical relations between elements in a sentence. As Berkoff (1982: 12) notes in his study, the use of prepositions in their local meanings seemed to be easier for learners. However, Finnish differs considerably from Germanic languages in its use and even in its grasp of local prepositions in connection with certain verbs.

Finnish has an interesting system of expressing spatial relations. These relations involve the contrast between 'external' and 'internal' location. The case endings with *l* are 'external' in that they are used to express location 'on top of something': thus the ending *-lla* or *-llä* can be translated as *on*. The case endings with *s* seem to refer to something 'internal', roughly corresponding the idea of containment (Hakulinen and Karlsson 1979). Thus the endings with *-ssa* -or *-ssä* will be translated as *in* (cf. above). Finnish uses this ending to express attachment, as well, for which English uses the preposition *on*. It was hypothesized that Finnish language learners would show a tendency to accept erroneous English structures with the preposition *in*.

Similar constructions were also used in the Swedish sentences included in the experiment. Two contructions contained expressions that were supported by learners' L1. Finally, Finnish sentences with a similar use of *-ssa* or *-ssä* were also included, although not directly matched, in the Finnish language set. Eight such expressions with their grammatically acceptable counterparts were included in each set.

Mixed with the eight pairs of test sentences were also four correct-incorrect pairs of filler sentences that functioned as controls. The English control sentences served as a baseline since they were of a kind that had received relatively fast and unanimous judgments in Experiment 1. Originally, the Finnish and Swedish controls were intended to be used in the same way; however, as it turned out, some of the Finnish controls in particular turned out be quite difficult. For this reason, only four of them were used to provide a baseline for the Finnish sentences.

9.5. Research hypotheses

The following hypotheses were made concerning the relationship among the variables.

Hypothesis1. It was hypothesized that the NSs of English would be faster and more accurate than the learners when judging English sentences. It was also hypothesized that the Finns would be significantly faster in their judgments of Finnish L1 than English L2 or Swedish L3 since it was assumed that performance in the native language is largely automatized.

Hypothesis 2. It was hypothesized that the Finnish learners'judgements would reveal a tendency to rely on L1. It was hoped that the learners' judgments and judgment times shed some light on the processes involved in making decisions of this kind. This was all the more interesting because in this particular area of grammar, the learners' mother tongue, Finnish, appears to divide the conceptual space in a way different from English.

Hypothesis 3. It was hypothesized that there would be differences as to the length of the RT latencies discovered in grammatical acceptability tasks between individual language learners. It was in this area in particular where various cognitive variables were felt to be of explanatory value.

Hypothesis 4. Consequently, it was hypothesized that these differences could be related to factors such as field independence/dependence, verbal and non-verbal thinking style, and learners' ability to reason in L1. The learners' L2 communicative anxiety was also hypothesized to play a role. A pilot experiment by the present writer seemed to indicate that there exists a weak but significant correlation between an individual's score on EFT and her mean RT latency in grammatical acceptability tasks.It was tentatively hypothesized that the responses of the more field-independent subjects would be slower but they would be more accurate in foreign language sentences, as they are generally held to be more "analytical", and less influenced by context.

9.6. Procedure

The items were prepared with the help of an Amiga software package DePaint II. In their preparation, similar guidelines were adhered to as in Experiment 1. The reaction time experiment took place during a two-week period at the English Department of the University of Jyväskylä. All the sentences and instructions were presented to the subjects on an Amiga 1000 computer by using a software package called Event Organizer. This program, originally developed by Seppo Sneck for psychological research, allows the experimenter to vary the nature and time of exposure of stimuli; the program also records the subject's response and reaction time.

The sentences were shown in a quasi-random order: The subjects were always presented first with the incorrect structure, and the correct stucture was shown 4-6 sentences later. Before the beginning of each set of test sentences, the subjects were shown the instructions for each set, and then five practice items. The order of the languages was English, Swedish, and Finnish. All the stimuli were visible for 2000 msec. A cut-off value of 6000 msec was used to exclude prolonged RTs. The screen flashed red 3.5 sec before each sentence in order to prepare the subject for the appearance of the next item.

The subjects were asked to decide as quickly as possible whether the sentence was in their opinion acceptable or non-acceptable as 'normal, everyday English', and then push the button marked accordingly. The purpose of such instructions was to elicit as intuitive responses as possible under such circumstances.

To measure the subjects' field dependence and independence, the Embedded Figures Test was administered to each subject individually. The subjects score was the mean (in seconds) of his performance time for the twelve items (1A to 12H) used in the study. The test was administered according to the instructions given in the EFT Manual (Witkin et al. 1971). Each subject was also given the L1 verbal reasoning test used by the Finnish Department of Labor for volitional counseling. The test consisted of a short text in Finnish L1 with a number of blanks, which the subjects had to fill in with a suitable word. The scoring of the questionnaire was according to the guidelines provided by the Ministry, with one point given for each acceptable alternative. The individual tests took place within one month of the RT test.

9.7. Statistical analysis

For the purposes of a statistical analysis, the reaction time data were transferred to theJyväskylä University main frame computer. The statistical program package SPSS was used for the analysis of the data. Descriptive statistics, frequencies, correlations, t-tests and analyses of variance were performed.

9.8 Results and Discussion

9.8.1 Hypothesis 1

It was hypothesized that the NSs of English would be faster and more accurate than the Finns. Figure 22 below shows the the descriptive statistics of the speed and accuracy of the

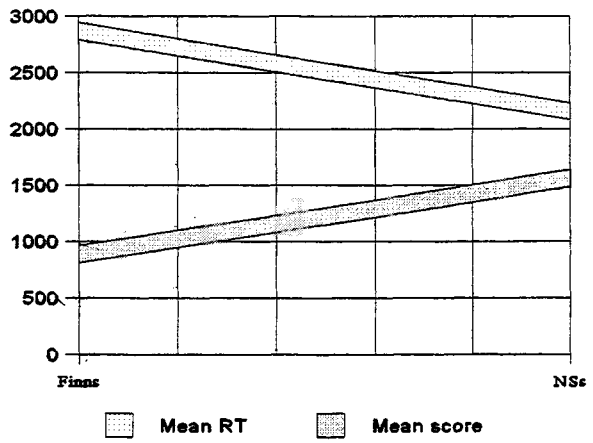


FIGURE 22 The mean RTs and scores for the target (preposition) sentences (k=16) for the Finns and NSs of English. (To fit the scale, the scores have been multiplied by 100.)

subjects' judgments in English. The difference between the means for the sentences turned out to be almost significant (df=6.21, T=2.41, p=0.05). This in itself was rather interesting. In Experiment 1, the NSs were significantly faster than the Finnish learners. Yet, the difference was still rather great. It is also quite likely that the small number of NS subjects influenced the results.

The NS of English were significantly more accurate in their judgments (their average raw score was 15.6 (sd = .55) out of 16, in contrast to the Finns, whose mean raw score was 8.9 (sd = 2.1). The difference in the means was highly significant (df=23.42, T=-13.07, p=0.000). Although the NS group was expected to be more accurate in their acceptability judgments, the result was nevertheless unexpected, since the Finnish subjects were all advanced students of English.

Figure 23 shows the descriptive statistics for the control sentences. The difference in the mean RTs was statistically not quite significant in the case of these sentences

(df=4.54,T=.62, p=0.562). However, it was significant for the mean scores (df=13.26, T=-4.01 p=.001), even though the difference in raw scores, 7.8 (sd = .45) for the NSs and 6.7 (sd = .91) for the Finns, appeared to be slight.

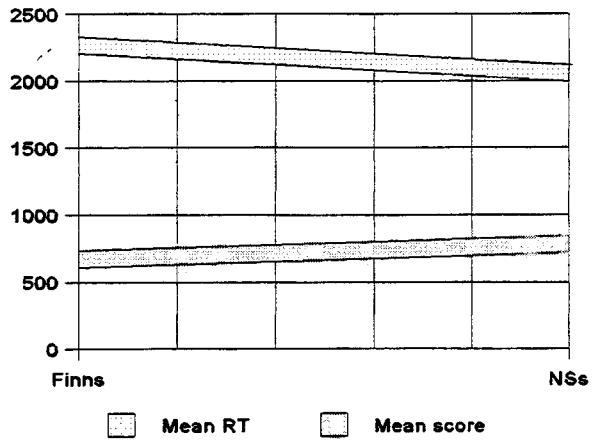


FIGURE 23 The mean RTs and scores for the control ('easy') sentences (k=8) for the Finns and NSs of English. (To fit the scale, the scores have been multiplied by 100.)

This may have partly been due to the obvious mixture of 'slow' and 'fast' processors within each group: the small sample (N=5) of the NSs group biased the results since two 'slow' processors were included. Even so, the Finns' performance was slower and more inaccurate than that of the NSs for both types of sentences.

Still, there was a difference for the Finns between the 'preposition' and 'easy' sentences: A paired t-test revealed that the difference between mean RTs was statistically significant (T=8.74, df=20, p=.000). This would seem to indicate that the judgment of these particular structures caused difficulty for the learners.

Figure 24 shows the results for the Finnish learners' RTs and the judgments on the target structures in English L2, Swedish L3 and Finnish L1. The MANOVA with the

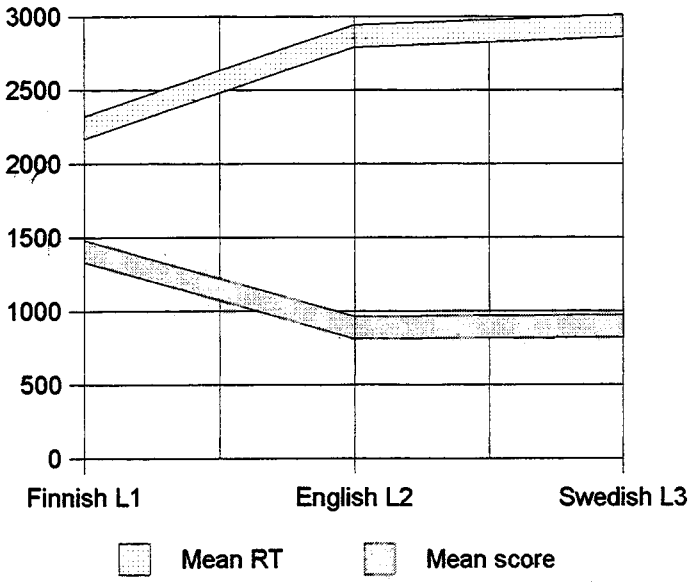


FIGURE 24 The mean RTs and scores of the Finns for the target (preposition) sentences (k=16) in all three languages. (To fit the scale, the scores have been multiplied by 100.)

the Wilks' multivariate test of significance was used to examine the statistical significance across the mean reaction times in the English, Swedish and Finnish language tasks. The effect of the language was extremely significant (Wilks' value=.21845, Exact F=33.98762, Hypothetical df= 2.00, Error df= 19.00, p=.000). The application of the paired t-test also showed that the difference in the RTs between the English and Swedish target sentences was not statistically significant (df=20, T=-.74, p=.470). Similarly, the difference between the mean scores was not statistically significant (df=20, T=.34, p=.735).

However, there was a statistically significant difference between the mean RTs of the English and Swedish control sentences (df=20, T=-5.21, p=.000). Figure 25 shows the mean RTs for the control ('easy') sentences in all languages. The mean scores have not been included but for English the mean score was 6.7 (sd = .91), Swedish 5.1 (sd = .94), and Finnish 3.6 (sd = .22); only four control sentences were used for Finnish.

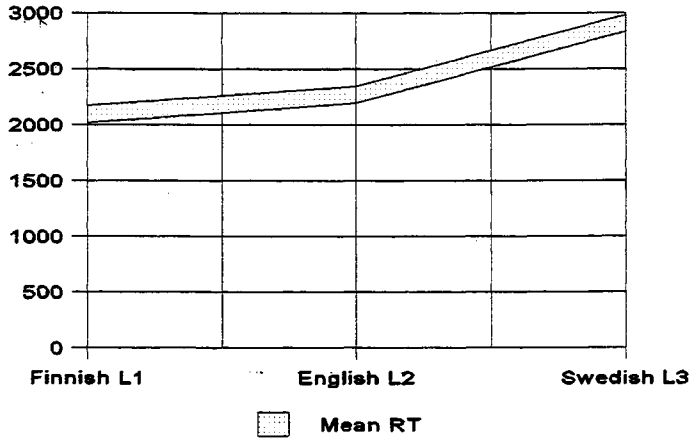


FIGURE 25 The mean RTs of the Finns for the control ('easy') sentences in all three languages.

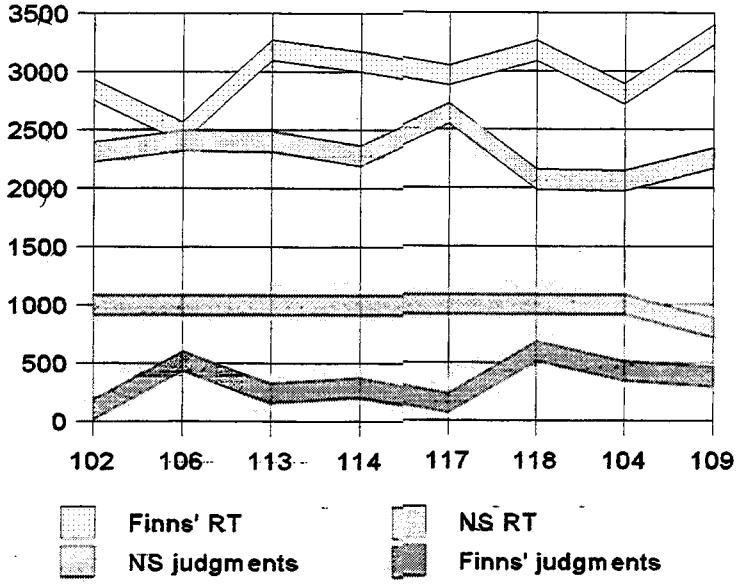
The Finns' accuracy in judging the English L2 structures was rather poor: the raw scores they received in English L2 were actually equal to those they had in Swedish L3. (There was a great deal of great variation among the learners. One Finn, although her scores were not included in the statistical analysis because of her long stay in an English-speaking country, actually judged all 16 sentences correctly.)

In all these cases, however, the NS judgments and judgment times show a pattern typical of fluent, skilled behavior. Whether they were NSs of English judging English or NSs of Finnish judging Finnish, the judgments were fast and accurate. Depending on the type of linguistic material to be judged, the learners showed a great deal of variation: they judged 'easy' areas of grammar almost as fast as the NSs; their accuracy, though, was not as high. The learners also judged the Swedish 'easy' sentences slower than the English 'easy' sentences.

9.8.2 Hypothesis 2

Figure 26 shows learner and NS performance on the erroneous target sentences According to Lehtonen and Sajavaara (1985), a fast acceptance of an ungrammatical sentence can be interpreted as transfer of more automatized processes from L1. There was one example (sentence 106) of this type in the data. Sentences 102, 104 and 117 also received judgment times that were comparatively low. The difference in the mean reaction time latencies

between the Finns and the NSs was almost significant only for sentence 104 (t= 2.32, df = 8.26, p =.048). In addition, there were two other deviant target sentences (109 and 118), for which the Finns' judgments were almost significantly slower. For all these sentences, the NSs spent a longer time than the average for this task; for example, sentence 106 **She thought Jack was in the phone* was judged by the NSs relatively slow (mean = 2311 msec, sd = 618 msec) in contrast to the Finns, who were not very slow (mean = 2841 msec,



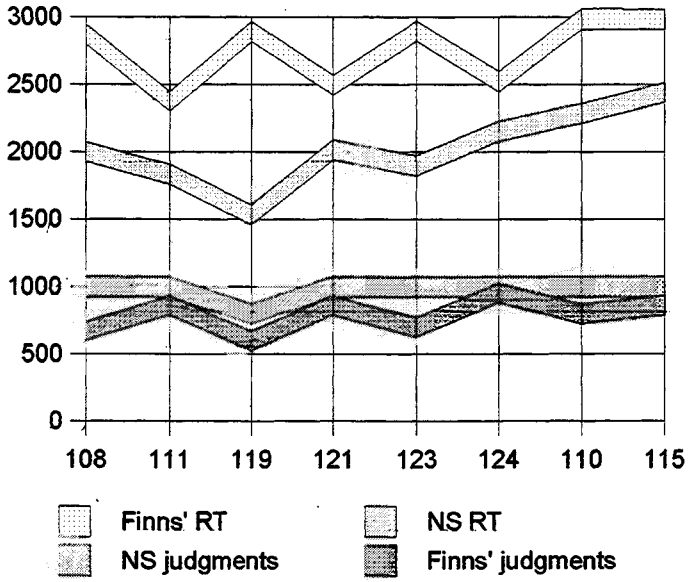
102 *Amy had a golden ring in her finger
106 * She thought Jack was in the phone
113 *Jim could see no handle in the door
114 *Ted thought his hair was in fire
117 *Both countries are now in war
118 *They wanted to sit in her table
104 *He saw his face at the mirror
109 *Ann saw a hat on the shop window

FIGURE 26 The speed and accuracy of judgments by the Finns and NSs for sentences in which the Finns were almost significantly slower. To fit the scale, the scores (as percentage points) were multiplied by 100, ie. 100% was transformed into 1000.

sd=554 msec). This lengthening of the RT was probably caused by the unusual meaning conveyed by the sentence, of which the NSs appear to have been more aware. Similarly, sentence 117 was processed relatively slowly by the NSs of English.

Sentences 104 **He saw his face at the mirror* and 109 **Ann saw a hat on the shop window* were incorrect counterparts of structures for which the preposition *in*, equivalent of the Finnish case ending *-ssa* or *-ssä*, would have actually been correct. To see whether the Finns' performance on sentences that conformed to their L1-based expectations (110 and 115) differed form their performance on other sentences, the correct sentences were compared with each other. Figure 27 shows all learner and NS performance for all the correct target sentences. It appears that both the Finns and the NSs of English required slightly more time to judge sentences 110 and 115, although both groups were quite accurate. On the other hand, sentences 108, 119, and 123 obviously caused some problems for the learners, since they were judged comparatively slowly and less accurately.

The correct sentences received much more accurate judgments. Again, the Finns' judgment times were longer than those of the NSs. The Finns were slower than the NSs to a significant degree in judging the correct sentences 119 *There was no handle on the door* (t= 3.79, df=19.44, p= .001) and 123 *She likes to sit at Joe's table* (t= 3.02, df= 17.32, p= .008). For 111 *He heard them talking on the phone*, the Finns were almost significantly slower. There were no significant differences between the RT means for the other sentences.



108 You have a new ring on your finger
111 He heard them talking on the phone
119 There was no handle on the door
121 France and Germany were at war
123 She likes to sit at Joe's table
124 The said our house was on fire
110 Jill noticed Bill in the mirror
115 She liked the shoes in the window

FIGURE 27 The speed and accuracy of judgments by the Finns and NSs on the correct target sentences. To fit the scale, the scores (as percentage points) were multiplied by 100, ie. 100% was transformed into 1000.

In general, learner performance on erroneous sentences appeared to differ in certain respects from that of the NSs. Figure 28 shows the average speed of judgment on both correct and incorrect sentences for the Finns and NSs of English. A t-test revealed that the difference between the subject groups was almost significant for the correct sentences (t=2.58, df=6.71, p=.038), but not for the incorrect sentences, although the difference was still quite large (t=2.15, df=6.19, p=.074). On the whole, the erroneous sentences received longer RTs from both groups.

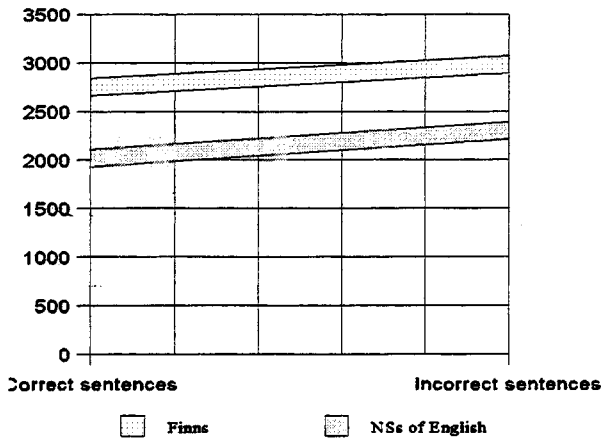


FIGURE 28 The average speed of judgments for correct and incorrect English L2 target sentences by the Finns and NSs of English.

However, as mentioned above, the Finns were remarkably inaccurate in their judgments. Figure 29 shows the accuracy of learner and NS judgments. For both types of sentences, the Finns were less accurate than the NSs to an extremely significant degree ($t=-4.66$, $df=19.90$, $p=.000$ for the correct sentences; $t=-13.42$, $df=22.18$, $p=.000$ for the incorrect sentences.

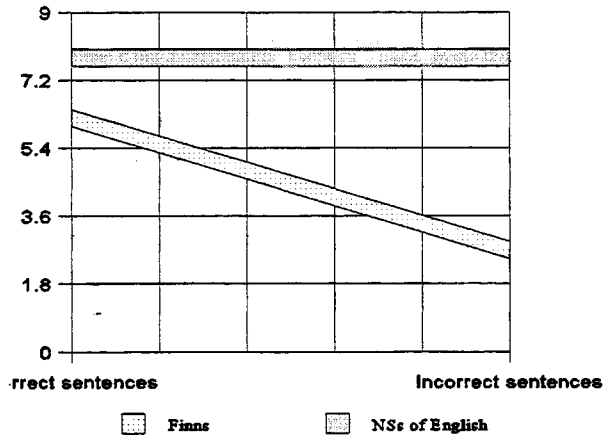


FIGURE 29 The accuracy of judgments as mean scores for correct and incorrect English L2 target sentences by the Finns and NSs of English.

For example, sentence 102 **Amy had a golden ring in her finger* was accepted by as many as 19 out of 21 Finns. Seven Finns rejected the correct sentence 108 *You have a new ring on your finger*. Of the subjects who rejected the correct sentence, six (Ss 1,4,6,11,12, and 21) (except 17, who rejected both) also accepted the erroneous sentence. The inevitable conclusion is that the interlanguage of these learners very likely contained a non-target-like knowledge structure or processing routine from which they did not vary at least during this experiment. Moreover, a similar tendency appeared in Swedish. Figure 30 shows the speed and accuracy of the Finns' judgment of the Swedish sentences containing

218 *Hon hade en ny ring i fingret
224 Erik hade en ring på fingret

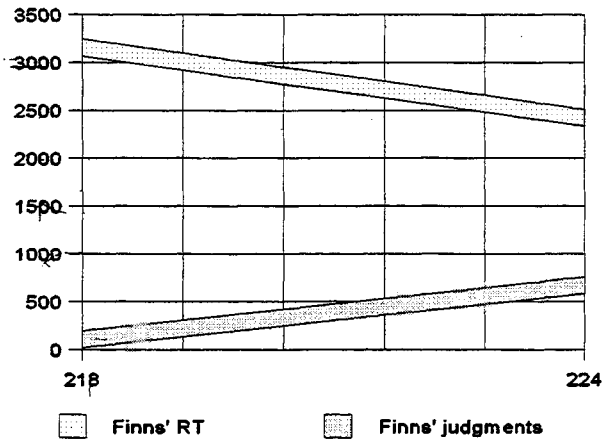


FIGURE 30 The speed and accuracy of judgments by the Finns and NSs for sentences in which the Finns' L1 contains a similar expression. To fit the scale, the scores (as percentage points) were multiplied by 100, ie. 100% would be transformed into 1000.

a similar expression to 102 **Amy had a golden ring in her finger* (the Swedish preposition *i* is the equivalent of *in*, while the preposition *på* is often translated by *on*). The results seem to indicate that this particular structure presented the Finnish learners with serious problems. One reason for this may have been its infrequency. In foreign language learning contexts, it is not likely that learners are frequently exposed to or have to use such spatial relationships as expressed in sentence 108 *You have a new ring on your finger*, or even in sentence 109 *There was no handle on the door*.

Because of the relative unfamiliarity of these constructions, the learners appeared to rely heavily on their L1-based knowledge or processing mechanisms. It may well be that during the second language learning process the learners have acquired pieces of procedural knowledge that consist of IF-THEN pairs containing instructions on how to find equivalents for expressions in one language within another language. For example, a good case could be made on the basis of these findings that at least these learners approached the task with the specific set of procedures in mind, such as: IF there is *-ssa* or *-ssä* in Finnish, THEN use *in* in English; or vice versa: IF there is *in* in English, THEN use *-ssa or -ssä* in Finnish. It is likely that formal language instruction encourages the formation of such strings of knowledge, and that in time they may become so routinized that the learner has become quite automatic in their execution and is also no longer aware of them. Whether they become learners' implicit interlanguage knowledge is difficult to tell on the basis of these experiments; however, they seem to direct the access to linguistic information during performance.

To sum up, there was an extremely significant tendency for the Finns to accept the sort of erroneous prepositional constructions that were included in this experiment. In most cases these sentences also received longer judgment times. It would thus appear that most of the structures included in the experiment type were not automatized to the same degree as they were among the NSs, since for the most part, the learners were unable to judge prepositional structures such as these quite fluently and accurately. Interaction between L1 and L2 (and even L3) knowledge and interlanguage system-internal factors may have caused these areas of language knowledge remain relatively 'fuzzy' in the minds of the learners.

9.8.3 Hypothesis 3

On the basis of the results of Experiment 1, the individual learners' RT means were analyzed to see whether the subjects' mean RTs in different languages were correlated with each other. Table 6 below shows the results of the Pearson product-correlations calculated among the judgments times.

TABLE 6 Pearson product-moment correlations among the English L2 , Swedish L3 and Finnish L1 target sentences and the English L2, Swedish L3 and Finnish L1 control sentences. ** = p≤ .01

	English	Swedish	Finnish	English Control	Swedish Control	Finnish Control
English	-	.82**	.69**	.90**	.68**	.69**
Swedish		-	.73**	.70**	.86**	.67**
Finnish			-	.68**	.79**	.67**
English Control				-	.66**	.57**
Swedish Control					-	.56**
Finnish Control						

The coefficients indicate that there is, indeed, a strong correlation among the speed of judgments in various languages, the relationship being the strongest between English L2 and Swedish L3. The correlation between English L2 and Finnish L1 is significant as well, although not quite as powerful. This would seem to indicate that those individuals who were slow in performing this kind of task in one language were slow also when they were asked to make judgments on the other languages.

The difference in the means between English and Swedish for the Finns was statistically not significant (t=-.74, df=20, p=.470); the subjects' proficiency in Swedish was not checked before the experiment but it turned out that only two of them studied Swedish actively at the university. Still, this result was not unexpected because both task groups contained similar items. Moreover, the English sentences were always presented first, so they may have activated the relevant knowledge in the language learners. Some slight support for this was provided by the fact that there was a weak correlation between the accuracy scores for both the English and the Swedish sentences. Again, there was no statistically significant difference in the two mean scores (t=.34, df=20, p=.735); as a matter of fact, the mean score for Swedish was slightly higher than that of English.

Similarly, there was a rather strong positive correlation between Finnish L1 and Swedish. The mean reaction times for English and Swedish were approximately 800-900 msec longer, however, than for Finnish, which implies an average increase in the processing load in foreign language tasks.

The correlation coefficients between the scores for the target and control sentences were not significant, nor did the control sentences correlate significantly among each other. The scores for the English and Swedish target sentences appered to be positively related (r= .44 with p ≤ .05). The mean judgment time for the Finnish control sentences had a significant negative correlation (r = -.56) with the accuracy with which the sentences were judged. In other words, the faster the judgments, the more accurate they were. This is a feature generally held typical of automatic skilled performance. The correlation coefficients between the scores and RTs are shown in Table 7 below.

TABLE 7 Pearson product-moment correlations among the RTs and scores of English L2 , Swedish L3 and Finnish L1 target (KEng, KSwe, KFin) and control (NKEng, NKSwe, NKFin) sentences. ** = p≤ .01

	KEng	KSwe	KFin	NKEng	NKSwe	NKFin
English RT	n.s.	n.s.	n.s.	-.47*	n.s.	-45*
Swedish RT	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Finnish RT	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
English Control RT	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Swedish Control RT	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Finnish Control RT	n.s.	n.s.	n.s.	n.s.	.44*	-.55*

The subjects were also ranked according to their RT means for both the English L2 target and control sentences. Figure 31 shows the Finns' mean RTs for the target sentences. It turned out that the relative order of the language learners remained constant: For the English L2 targets, Ss 1,5,17,19,15, and 3 (in this order) had the slowest mean RTs, ranging from 4056 msec to 3331 msec. For the English L2 control sentences, Ss 5,19,1,15, 17, and 10 had the longest reaction times, ranging from 3112 msec to 2557 msec. Thus, the same

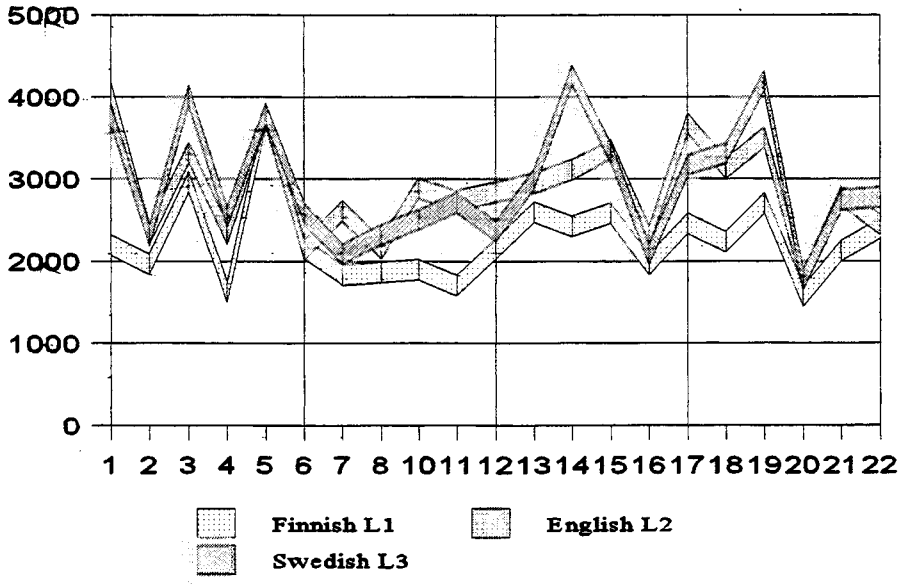


FIGURE 31 The Finnish subjects' mean RTs for each of the target languages.

five subjects were among the slowest performers in English regardless of the nature of the stimuli. On the other hand, the fastest performers in the English L2 targets were Ss 20,6,8,16, and 2, ranging from 1765 to 2329 msec. For the control sentences, the fastest were Ss 20,6,22,2, and 8, ranging from 1584 to 1933 msec. Only Ss 22 and 16 had changed places.

On the basis of these results, two *post hoc* subject groups, 'Slow' and 'Fast' processors, were formed, the 'Slow' including Ss 1,5,17,15 and 19, the 'Fast' Ss 2,6,8,16 and 20. Figure 32 shows the mean judgment times for these two groups in all three languages.

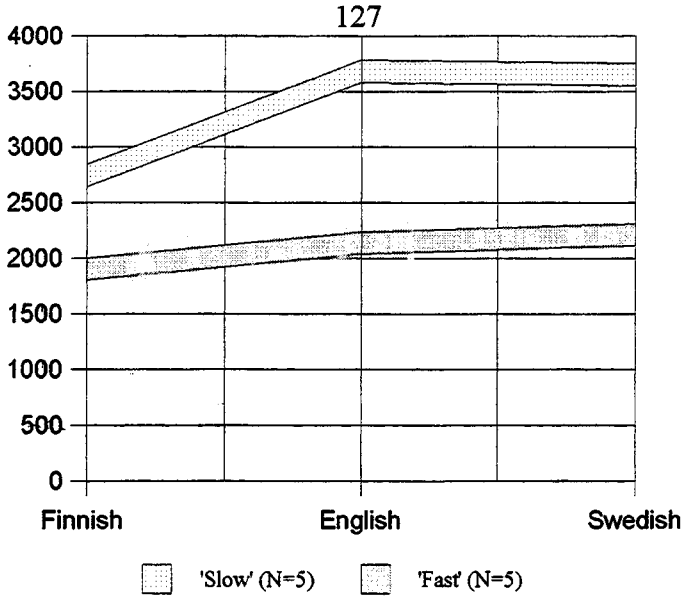


FIGURE 32 The judgment times of the 'Slow' (N=5) and 'Fast' (N=5) Finnish L1 learners of English in Finnish, English and Swedish target sentences.

The ANOVA revealed that, in addition to the English sentences, there was a significant difference (at the .05 level) between 'Slow' and 'Fast' subjects as to the mean RTs for the Finnish L1 target sentences (df=4.92, t=2.94, p=.033) and control sentences (df=7.03, t=2.56, p=.037), and the Swedish L3 target sentences (df= 7.43, t=6.30, p= .000) and control sentences (df= 5.96, t= 3.50, p=.013).

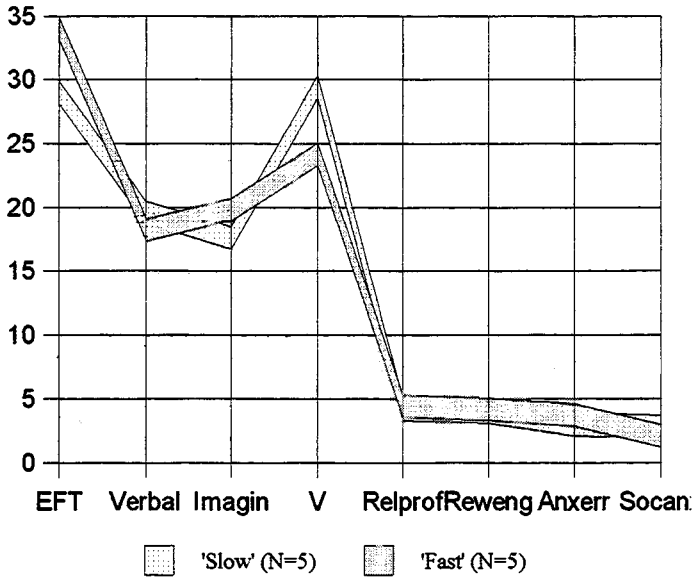


FIGURE 33 The 'Slow' and 'Fast' processors in relation to the cognitive variables of EFT (field dependence/independence), Verbal (verbal style of thinking), Imagin (the use of imagery in thinking), V (verbal reasoning in L1), Relprof (reliance on one's language proficiency), Reweng (reward received in English classes), Anxerr (anxiety about errors) and Socanx (social anxiety). The higher the EFT score, the more field-dependent the subjects are.

How these two groups differed in relation to the various cognitive variables is shown in Figure 33. There was a not quite significant difference between these subjects as to the factor PF4 ('absence of reading difficulties') (t=2.21,df= 7.69,p=.059) and IMAGIN ('use of imagery') (df= 7.26, t=-2.20, p=.062) based on the IDQ. The direction was opposite to what might have been expected: 'Slow' judges scored higher than 'Fast' judges when the number of such statements in the IDQ were counted in which the subjects described themselves as good readers, or as ones to whom reading was no problem. Similarly, 'Fast' judges scored higher than 'Slow' judges as to the use of imagery in thinking. This might indicate that the fast judges rely more on holistic strategies in sentence processing. Interestingly, Levelt et al. (1977) propose that concrete, high-imagery language items would receive faster RTs than more abstract ones. These results seem to suggest that the subjects who tended to be more visually oriented and inclined to be more holistic in their thinking might be faster in problem-solving tasks like this.

9.8.4 Hypothesis 4

After the calculation of the Pearson product-moment correlations among the target sentence judgments and judgment times and the various cognitive variables, very few strong correlations were found, as is shown in Table 8. The two dependent variables that were of the greatest interest, namely ENG (=mean reaction time latency for English target sentences) and KENG (=mean score in the judgments on the English target sentences), proved to have few links to cognitive variables. The results of the correlation analyses will be discussed according to the variable group.

Speed

There were six RT variables in the experiment: ENG, SWE, and FIN were the mean RTs for the target prepositional structures, while NENG, NSWE and NFIN were the mean RTS for the 'easy' control sentences. Their relationship with each other is shown in Table 5. As was mentioned above, these variables were highly related to each other. Of the target sentences, SWE had a negative correlation with RELPROF (r=-45*) and ANXERR (r=-44*). FIN had a significant negative correlation with RELPROF (r=-.46*) and ANXERR (r=-.44*) as well. ENG did not correlate with any of the cognitive variables. In addition, NFIN was positively correlated with REWENG (r=.46*) and NENG was correlated with V (r=.44*). This was rather interesting as it seemed to suggest that higher RTs on the English 'easy' sentences were related to higher scores in the L1 verbal reasoning test, ie. the better the subjects' L1 verbal reasoning was, the higher their RTs were on at least some of the L2 sentences. Moreover, greater reward in English classes seemed to be related to higher RTs in Finnish. No other control RT variables were related to any of the cognitive variables.

Accuracy

There were six accuracy variables in the experiment: KENG, KSWE, and KFIN were the mean accuracy scores for the target prepositional structures, while NKENG, NKSWE and NKFIN were the mean accuracy scores for the 'easy' control sentences. Of these, KSWE was significantly related to PF4 ('absence of reading difficulties') (r=.46*). NKENG had a relatively strong positive correlation with RELPROF (r=.60**)

TABLE 9 Correlation coefficients among the cognitive variables. *=p=<.05,**=p=<.01.

	Wit	Relprof	Reweng	Anxerr	Socanx	PF1	PF2	PF3	PF4	PF5	PF6	PF7	Verbal	Imagin	V
Wit	-	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	ns.	-.44*
Relprof		-	.49*	.46*	.59**	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	.56*	n.s.	n.s.
Reweng			-	n.s.	n.s.	n.s.	.44*	.44*	n.s.	.51*	n.s.	n.s.	n.s.	.51*	n.s.
Anxerr				-	.50*	.58**	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Socanx					-	.76**	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	.64**	n.s.	n.s.
PF1						-	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	.72**	n.s.	n.s.
PF2							-	.44*	n.s.	n.s.	n.s.	n.s.	n.s.	.90*	n.s.
PF3								-	n.s.	n.s.	n.s.	n.s.	.60*	n.s.	n.s.
PF4									-	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
PF5										-	n.s.	n.s.	n.s.	.56*	n.s.
PF6											-	n.s.	n.s.	.47*	n.s.
PF7												-	n.s.	n.s.	n.s.
Verbal													-	n.s.	n.s.
Imagin														-	n.s.
V															-

TABLE 8 Correlation coefficients between the language judgment and cognitive variables. *=p=<.05,**=p=<.01.

	Eng	Swe	Fin	KEng	KSwe	KFin	NEng	NSwe	NFin	NKEng	NKSwe	NKFin
Wit	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Relprof	n.s.	-.45*	-.46*	n.s	n.s.	n.s.	n.s.	n.s.	n.s.	.60**	n.s.	n.s.
Reweng	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	.46*	n.s.	n.s.	n.s.
Anxerr	n.s.	-.44*	-.44*	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Socanx	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
PF1	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
PF2	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
PF3	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
PF4	n.s.	n.s.	n.s.	n.s.	.46*	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
PF5	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
PF6	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s
PF7	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Verbal	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Imagin	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
V	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	.44*	n.s.	n.s.	n.s.	n.s.	n.s.
Sweprof	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.

Field-dependence/independence

The variable WIT, which stood for the field-dependence/independence dimension of cognitive style, had a weak, yet significant, negative correlation with V (r=-.44*). It did not correlate with any of the other variables.

L2 communicative anxiety

There were four variables in the experiment that were based on the communicative anxiety questionnaire: RELPROF, REWENG, ANXERR, and SOCANX. Their correlations with the RT and accuracy variables have been discussed above.

RELPROF was positively related to REWENG (r=.49*) and ANXERR (r=*.46), and even more strongly to SOCANX (r=.59**). Because of the scoring, the results must be interpreted as follows: the more reliance the subjects reported as having on their L2 proficiency, the more reward they had reported having received in English classes and the less anxiety about errors, and even less about social situations.

REWENG was positively related to PF2 'habitual use of imagery' (r=.44*), PF3 'concern with the non-deviant use of the words' (r=.44*), PF5 'use of images to solve problems' (r=.51*), and IMAGIN, the score on the imaginal scale as a whole (r=.51*). It appeared that the subjects who reported as having received greater reward in English classes tended to use more imagery in their thinking. They also appeared to more concerned with the correct use of the words.

ANXERR was positively related to SOCANX (r=.50*) and PF1 'good verbal ability' (r=.58**). In other words, the less anxiety the subjects felt about errors, the less anxiety they felt about social situations. Moreover, they also appeared to have good verbal ability.

SOCANX was even more strongly related to PF1 'good verbal ability' (r=.76**). Thus, the subjects who reported to feel less anxiety about social situations on the CAQ also reported themselves as having good verbal ability on the IDQ. SOCANX also correlated relatively strongly with VERBAL, the score on the verbal scale as a whole (r=.64**), which was not surprising since PF1 and VERBAL turned out to be strongly correlated.

Verbal and non-verbal thinking styles

There were nine variables formed on the basis of the IDQ: VERBAL (the nominal scores on the original verbal scale), IMAGIN (the nominal scores on the original imaginal scale), PF1 'good verbal ability', PF2 'habitual use of imagery', PF3 'concern with the nondeviant use of words', PF4 'reading difficulties', or rather absence of them as this variable was counted as verbal in content, PF5 'use of images to solve problems', PF6 'vividness of dreams, day dreams,and imagination', and PF7 'verbal thinking'.

VERBAL correlated highly with RELPROF, SOCANX, PF1 (r=.72**) and PF3 (r=.60*). IMAGIN was positively related to REWENG, PF5 (r=.56*), PF6 (*.47), and even more strongly with PF2 (r=.90**). In other words, 81% of the variance in the IMAGIN could be explained by this single factor. PF2 and PF3 also appeared to be related (r=.44*).

It is to be noted that for the most part the correlation coefficients were quite low, even when they were significant. This affects their reliability considerably. In all likelihood, they best serve as indications of potential links rather than of any firm evidence. On the other hand, certain general directions seem to appear. Thus, the results obtained on the affective variables concerning L2 communicative apprehension seem to point to some kind of relation with the grammaticality judgment task.

9.8.5 Effect of language profiency

Since the Finns' level of English L2 language proficiency was not measured, it is not possible to look for its effect on the learner performance on English, except by comparing it to the performance of the NSs. As became evident in Chapter 9.8.1, the NSs were significantly faster and more accurate in this task, which can only be explained by their greater proficiency in English. All in all, the NSs behavior was typical of skilled, fluent performers.

It is possible to look at the effect of language proficiency also from another perspective. Comparison of the subjects' performance on correct sentences with their performance on incorrect sentences within each language shows that the subjects responded to incorrect L1 sentences differently than to incorrect L2 sentences. Figure 34 shows the performance of the NSs of English and the NSs of Finnish on their L1s, respectively.

The profiles are remarkably similar. For both the Finns and NSs of English, there was no statistically significant difference in the accuracy of their judgments between correct and incorrect L1 sentences. As to the mean RTs for these two types of sentences, the difference was not statistically significant, although it came close to it, in particular for the NSs of English: for the Finns, t= -1.98, df=20, p=.062; for the NSs of English, t=-3.05, df=4, p=.038.

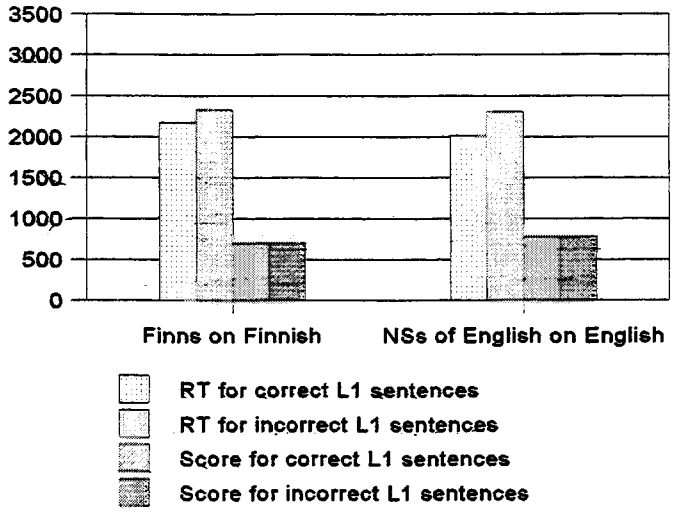


FIGURE 34 The subjects' performance on the correct and incorrect L1 sentences.

The contrast to the learner performance on the correct and incorrect L2 sentences was remarkable. Figure 35 shows the mean RTs and the scores by the Finns for both types of English L2 sentences.

There are several explanations to the observed difference between the judgment times of the correct and incorrect L2 sentences (cf. Hedgcock 1993) First, it might be due to the methodology: after all, the subjects were first shown the incorrect sentences. This might have caused a priming effect, even if care was taken in the experiment to ensure that the

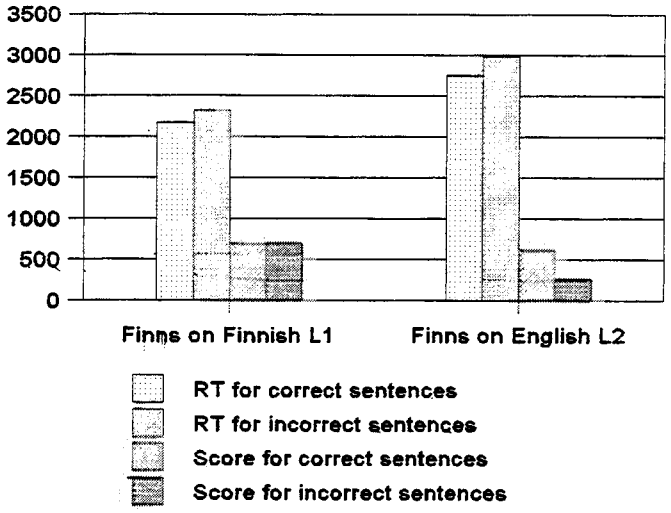


FIGURE 35 The Finns' performance on the correct and incorrect Finnish L1 and English L2 sentences.

incorrect sentences and their correct counterparts never followed each other. Still, some part of the information attended to during task performance could have caused an activation of the relevant knowledge in long-term memory. The fact that the judgment times of the incorrect L1 sentences did not differ significantly from those of the correct ones speaks against this explanation. However, this does not in itself mean that these RTs were not affected as there is no way of knowing whether the incorrect L1 sentences might have been judged faster.

It is also possible that the faster and more accurate responses to the correct sentences can have been caused by the retrieval from memory of similar instances: of course, it is not possible that exactly the same sentences were retrieved; yet it is possible that their frames were. This would speak for Logan's view of memory-based automaticity. This would mean, though, that correct and incorrect sentences would not undergo similar processing, that incorrect sentences would more likely be subjected to algorithm-based nonautomatic processes, or controlled processes in the sense Schneider and Detweiler (1988) use the term, that is, unless there were something in the sentences that would lead to fast acceptance, such as familiarity with the form because it occurs in one's interlanguage.

These findings seem to indicate that this type of linguistic task is sensitive to the level of subjects' language proficiency. Language proficiency appears to affect both the speed with which these judgments are made and their accuracy. The identification of erroneous L2 sentences is particularly problematic. It is even more likely that the effect that language proficiency appears to have on these results could be explained by differences in the underlying automatic procedural skills, ie. fluency.

9.9. Conclusion

In Experiment 2, an attempt was made to find out whether there was a relationhip between the speed and accuracy of language learners' acceptability judgments and a group of cognitive and affective variables that included language learners' cognitive style of field-

dependence/independence, L2 communicative anxiety and L1 verbal reasoning. In addition, learner and NS performance was compared.

No significant correlations were found between field-dependence/independence and the speed and accuracy of learner judgments. There were a few fairly significant correlations between the affective variables of L2 communicative anxiety and the speed and accuracy of the judgments, indicating that such affective factors as anxiety about errors and reliance on language proficiency might influence the speed with which tasks of this kind were performed. This relationship manifested itself only in Finnish L1 and Swedish L2, however. Verbal and nonverbal thinking style and L1 verbal reasoning appeared to have no relationhip with performance on these tasks, either. All in all, the correlations were not strong, no stronger than those that have been obtained by others studying field-dependence/ independence or other cognitive and affective variables. The sheer complexity of the human mind makes it unlikely that any clear and straightforward relationships could be established among constructs that are not yet very well understood or well defined.

Some interesting results were obtained, though, when the relationship between the cognitive and affective variables was analyzed. It appeared that good verbal ability was related to a greater degree of confidence in social situations and a lesser degree of anxiety about errors.

The performance of the Finnish L1 learners of English differed significantly from that of the NSs. Not only were the Finns slower, but they were also less accurate. This was considered to be the result of the different levels of language proficiency. The subjects' performance on the correct and incorrect L1 sentences differed, as well. The difference was especially noticeable on the L2 sentences, as the incorrect L2 sentences led more frequently to error and were judged more slowly.

From the methodological point of view, these results are encouraging for several reasons: since it appears that acceptability judgments like these are capable of differentiating between NSs and learners of a language, it is quite likely that as a task type they are not quite as unreliable and useless as has been suggested by their harshest critics. A great care has to be excercised in the selection and preparation of the sentences, however, although the same criteria, of course, hold for any kind of tool used to gather language data. The results of this experiment are affected by some discrepancies between the L1, L2 and L3 items: for example, some items turned out to be ambiguous (especially in L1). The frequency of the items could not be controlled for. The Swedish items were not properly commesurate with the English items either.

In Experiment 3, an attempt was made to rectify some of the defects in the design of Experiment 2. The focus will be on the relationship between language learners' L2 profiency and their performance in the grammatical acceptability task.

10 Experiment 3

10.1 Research questions

On the basis previous findings, it was assumed that (see also Lehtonen 1989; Lehtonen and Sajavaara 1983; Sajavaara and Lehtonen 1989; Alanen 1987,1991; see also Bley-Vroman and Masterson (1989), who used reaction time as a supplement to a sentence-matching task):

1. Language learners have longer reaction times than native speakers (NSs) of the language in question;
2. Language learners' judgments tend to be less accurate than those of NSs;
3. There seems to be some sort of systematic variation as to the reaction time latencies.

Such findings raise the question of performance in such tasks somehow being related to the language proficiency of the subjects. To study the relationship of language learners performance in judging the grammatical acceptability of English L2 sentences to their English proficiency, the following series of experiments was set up. The eventual goal was to find answers to the following questions:

1. What is the relationship of English learners' grammatical acceptability judgments and the speed with which they make such judgments to their language proficiency?
2. What is the nature of this type of language task? What kind of factors are involved?
3. What does this type of language task tell us about the nature of language proficiency and linguistic knowledge?

10.2 Subjects

Sixty-six Finnish L1 first- and second-year students of English participated in the study. The subjects were all volunteers. Of sixty-six subjects, sixty-one took part in all tests. Twelve of the subjects were male, fifty-four female. Similarly, forty-nine of the subjects were English L2 majors while twelve had English L2 as a minor subject. (The difference lies mainly in the application process since both groups have to pass one and the same test to be able to enter the program.) Thirty-four of the subjects were first-year students and twenty-four second-year students of English while four subjects had started their studies three or more years before the experiment. For statistical purposes the latter two groups were handled as one. The average age of the subjects was 21.6 years while they had spent approximately ten years studying English in formal language learning contexts. They had also studied a number of other languages (including their Finnish L1) and could thus be considered as experienced language learners. On an average, they had spent 3.8 months abroad in some English-speaking coutry (see Table 10 below).

TABLE 10 The average age (in years), the average number of years of English study at school and languages studied (including Finnish L1), and the time spent in an English-speaking counry (in months).

	Mean	SD	
Age (N=66)	21.6	2.61	years
Years of English study at school (N=61)	10.0	.71	years
Number of languages studied (N=61)	3.7	.81	languages
Time of residence in an English-speaking country (N=61)	3.8	4.44	months

There was no great amount of variation among the language learners except as to how long they had spent in an English-language country. An initial assumption was made that the length of residence would not be related to the speed and accuracy of the judgments; after all, no such pattern had been revealed in Experiment 1.

What the subjects' English and Finnish grades had been in their high-school diploma and how they presently saw their skills in listening, speaking, reading and writing English are shown in Table 11 below. The highest possible grade is ten while the best possible assessment was five ("extremely good"). The subjects' English and Finnish grades were very good, and while the subjects thought their listening and reading skills were very good, they rated their productive skills slightly lower. All in all, the subjects were a very select group of people who had been very good at English in school.

TABLE 11 The subjects' English and Finnish grades in the high-school diploma (on the scale of 4 - 10) and their self-assessments (on the scale of 1-5) in listening, speaking, reading and writing skills in English.

	Mean	SD	Out of possible
English Grade in the high-school diploma	9.7	.51	10
Finnish Grade in the high-school diploma	8.9	.74	10
Self-Assessment on English L2 listening	4.6	.52	5
Self-Assessment on English L2 speaking	3.9	.78	5
Self-Assessment on English L2 reading	4.4	.69	5
Self-Assessment on English L2 writing	4.1	.68	5

10.3 Design of the experiment

As in the previous two experiments, there were two basic assumptions underlying this experiment: First, judgments of grammatical acceptability are seen as instantations of L2 learners' use of language knowledge, ie. linguistic performance; and second, the speed with with the task is carried out is seen to reflect the degree of automization of language knowledge.

The design of the experiment is typical of within-subjects multivariate research: The subjects were tested on a number of variables chosen on the basis of previous research to find out whether there is a relationship (correlation) between them and the speed and accuracy of grammaticality judgments.

First, it was of considerable interest to compare the subjects' performance in L2 with their performance in L1. Therefore, the subjects were also asked to judge a set of Finnish L1 sentences containing comparable errors. The subjects' simple reaction times were also tested to see whether the sheer speed of reactions might influence their response times.

The second variable selected was L2 proficiency. This is related to the larger issue of the validity of timed grammaticality judgments. It is to be noted that L2 proficiency was operationalized by using an old university entrance exam minus the listening comprehension section. This particular instrument was chosen because it was considered important to use an instrument capable of differentiating between subjects who were advanced, if not very advanced learners of English and whose English was, indeed, very good. No claims are made in this study as to this being the abolutely proper way of measuring L2 proficiency. Measuring L2 proficiency, let alone testing it, is a complicated issue. To avoid any misunderstandings, the examination is described in more detail below, so that it will be possible to see exactly what was measured.

Another variable thought to be worth studying was L2 communicative anxiety: Based on a previous experiment, it was considered important to find out whether the performance in this type of task was related to the anxiety that learners feel about using L2 and especially making errors in L2. Again, this is remotely connected to the larger issue of the validity of timed grammaticality judgments.

A fourth group of variables consisted of the repetition of the task in a written mode with an additional task of locating and correcting the errors. Would there be a difference in learner performance across task modalities? Did learners actually react to the errors they were thought to react to? This was seen as related to the issue of the reliability of grammaticality judgments.

Fifth, background data such as sex, age, years of English study, English grades, and self-assesments in the four subskills of English, etc. were collected from the subjects.

9.4. Tasks

There were eight tasks used in this experiment.

1. Judgments and judgment times for isolated English sentences (n=39);
2. Judgments and judgment times for comparable isolated Finnish sentences (n=22);

The task itself was fairly simple: just as in the previous two experiments, the subjects were asked to judge ('Yes' or 'No') the grammatical acceptability of sentences containing errors

in syntax, morphology, or lexicon such as **He heard Jeff to come in* and **He left it to Jill's room*, or the corresponding correct sentences *She heard Mary go out* or *She left it in Jack's car*.

The tasks chosen for this experiment were used as a tool to examine L2 learners' ability to make judgments on the grammatical acceptability of correct and incorrect sentences containing errors for which learners might be tempted to rely on L1 knowledge and/or processes. The specific areas were chosen on the basis of the two previous experiments. All of the errors were of the kind Finnish L1 learners could be expected to produce or to accept, as they were based on error analysis studies of learner English produced by native speakers of Finnish, but not all of them were based on their L1. Some of them occur in the speech of English L2 learners ragrdless of their L1 background, such as errors in the *to* infinitive or 3rd person *-s* (see eg. Newport and Johnson 1989). It is worth pointing out that some of the latter sentences are very simple indeed since the errors were relatively elementary in nature.

3. Judgments of the same set of English sentences and correction of the potential error in writing;

4. Judgments of the same set of Finnish sentences and correction of the potential error in writing;

These tasks served two purposes: First, it was designed to tap into more explicit metalinguistic knowledge. Second, it provided a good opportunity to check whether the learners had actually reacted to the structure or item they were expected to.

5. Simple reaction time. The aim was to discover whether the individual differences in the RTs found in Experiments 1 and 2 could be explained by slower reactions in general.

6. English L2 language proficiency test, which included two grammar subtests, translation, reading comprehension, and vocabulary tests in a multiple-choice format. The two grammar tests and the translation test formed actually one single subtest called "grammatical structures", but since it consisted of three very different task types each forming its own specific section, it was felt that the three sections should be scored separately. In Grammar 1, the learners were asked to choose an alternative that best corresponds to the usage in standard English. For example,

Example 1 A tax declaration form is_____ for a small child to understand.
(a) too difficult a thing
(b) a too difficult thing
(c) too difficult thing
(d) too a difficult thing

In Grammar 2, the learners were asked which alternative was correct:

Example 2 (a) The money he was owed was to be paid for in cash.
(b) The money he was owed was to be paid in cash.
(c) The money he was owed was to pay in cash.
(d) The money he was owed was to pay for in cash.

They were also asked which of the sentences was different in meaning from the others:

Example 3
(a) Dreaming of his girlfriend, he realised, would not help.
(b) The dreaming of his girlfriend, he realised, would not help.
(c) To dream of his girlfriend, he realised, would not help.
(d) To be dreaming of his girlfriend, he realised, would not help.

In Translation, the students were given a Finnish sentence with an underlined structure and asked to choose the alternative that would best translate that part of the sentence into English.

Example 4
Hän hoiti lapsensa show-elämään jo kymmenvuotiaasta lähtien.
(a) since they were ten then
(b) as they were ten already
(c) when they were still only ten
(d) even while they were ten

In Reading Comprehension, the students were given four texts. After the first text, the students were given five sets of alternatives. The students had to choose which one of three alternatives (a), (b) or (c) corresponded to the meaning of the text; if none of them did, they could choose alternative (d). In the other three texts, adapted from various text books on English language and culture, the students were given sets of four alternative continuations of sentences of which they had to choose the one that best fitted the context. The students were adviced first to read through the entire text carefully, and try to decide the intention of the text as a whole.

In Vocabulary, the students were tested on their command of vocabulary and idioms. They were instructed to choose the alternative that best fits the gap in the sentence.

Example 5
Only prompt action by the pilot of the DC10 _____ disaster.
(a) averted
(b) distracted
(c) thrust off
(d) turned down

7. Questionnaires. The questionnaire was used to gather data on background variables such as age, sex, years of English study, number of languages studied, time of residence in an English-speaking country. It also included a L2 communicative anxiety questionnaire consisting of 22 Likert-type statements. The items were divided into the following factors: reward in English lessons (k=4); reliance on one's proficiency (k=6); anxiety caused by errors (k=6); and social anxiety (k=6). Four of the items used as part of reliance on proficiency were also used as learners' self-assessments on listening, speaking, reading and writing skills in English. The questionnaire was based on a much more extensive 119-item long questionnaire on L2 communicative anxiety (Lehtonen, Sajavaara and Manninen 1985), which was used in Experiment 2.

10.5 Research hypotheses

The following hypotheses were made about the relationship of the variables:

Hypothesis 1. The subjects would be faster and more accurate in judging L1 sentences than L2 sentences.

Hypothesis 2. There would be a positive correlation between English L2 and Finnish L1 judgment times;

Hypothesis 3. There would be a positive correlation between English judgment times and anxiety about L2 errors

Hypothesis 4. There would be a negative correlation between English L2 judgment time and English L2 proficiency test score;

Hypothesis 5. There would be a positive correlation between the accuracy of English L2 judgments and English L2 proficiency test score.

10.6 Procedure

The experiment was conducted in three separate phases. Phase 1 consisted of the reaction time experiment. It took place during a four-week period at the English Department of the University of Jyväskylä. All the sentences and instructions were presented to the subjects on an Amiga 1000 computer by using a software package Event Organizer, as in Experiment 2. The items were prepared with the help of an Amiga software package DePaint II. Similar guidelines were adhered to as in Experiments 1 and 2. As mentioned above, Event Organizer allows the experimenter to vary the nature and time of exposure of stimuli; the program also records the subject's response and reaction time. The subjects were tested individually.

The subjects were first given a five-item test to measure their simple reaction time. The subjects were asked to react by pushing a button on a box connected to a computer port when the screen changed color from light grey. The reaction times were recorded by the computer. The subjects were then tested by presenting them a set of correct and incorrect pairs of English L2 and Finnish L1 sentences. The sentences were presented in two blocks, one for each language. The order of presentation of the blocks was counterbalanced across the subjects so that 35 subjects responded to English sentences first while 31 subjects saw the Finnish sentences first. At the beginning of the block the subjects were given both oral and written instructions in Finnish.The written instructions appeared on the computer screen. After that, they were shown five practice items in both blocks. In addition, the first sentence in both blocks was always the same and was not part of the actual set of test sentences (see eg. Robinson and Ha 1993).

The entences were presented visually on a computer screen in a random order one at a time. The order was determined in advanced; if more than four correct or incorrect items followed each other, the order was broken to prevent response expectancies from developing. The sentences were visible for 2000 msec. The cut-off value was set at 6000 msec. The subjects responded by pushing a button on a box connected to a computer port. The response times (in msec) and responses were recorded automatically by the program.

Immediately after, subjects were given the same set of sentences in exactly the same order on 2.5" by 4" cards. The subjects were asked, first, to write down the response they

gave in the first test (by using letters *E* for *'Ei hyväksyttävä'* ' ("Not acceptable") and *H* for *'Hyväksyttävä'* ("Acceptable"). Secondly, they were asked to judge the sentence again and write down their present judgment by writing *K* for *'Kyllä hyväksyn'* ("Yes, I accept") if they accepted the sentence. If they did not accept the sentence, they were not required to write down a separate answer, although they could write down an *E* for *'En hyväksy'* ("Not acceptable") again if they wanted. Instead, they were asked to locate and correct the error they thought the sentence contained. The time subjects spent on each block was measured although the subjects were not told of this. The instructions were given both orally and in writing. Unknown to the subjects, the time the subjects spent on both English and Finnish set of sentences was recorded as well.

In Phase 2, the subjects were tested in larger groups. 2-4 weeks later each subject participated in the multiple-choice language proficiency test.

In Phase 3, the subjects were asked to fill in the questionnaire.

Out of originally 70 subjects who agreed to take part in the experiment, 66 completed the RT test. The RT data of two of the subjects were lost and two subjects were later discarded as their L1 was not Finnish. Of these 66, 61 returned the questionnaire.

10.7 Results and discussion

10.7.1 Hypothesis 1

It was hypothesized that the subjects would be faster and more accurate in their judgments of Finnish L1 sentences, The means and standard deviations in the grammaticality judgments of Finnish L1 and English L2 sentences are shown in Figure 36 and Table 12 below. The learners' mean SRT was 272 msec (SD = 66).

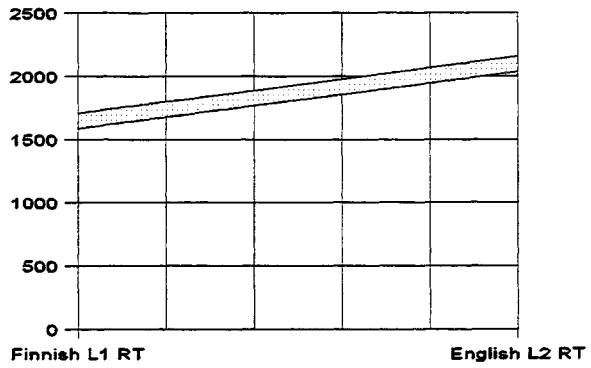


FIGURE 36 The speed of learner judgments in Finnish L1 and English L2 means and standard deviation (in msecs) (N=66).

A paired t-test revealed that the learners' mean RT in Finnish (mean = 1646, sd=309) was significantly faster than in English L2 (mean= 2097,sd=412). It also showed that the Finns were significantly more accurate in their Finnish judgments than English judgments. To compare the accuracy of Finnish L1 and English L2 judgments, the scores were first transformed into error scores. Table 3 shows the raw scores for learner judgments in the RT and written judgment tasks in Finnish L1 and English L2.

Variable	Mean	SD
English L2 Judgments (k=39)	29.6	3.1
Written L2 English Judgments (k=39)	34.5	2.7
Finnish L1 Judgments (k=22)	20.1	1.4
Written Finnish L1 Judgments (k=22)	21.0	0.9

TABLE 12 The mean scores and standard deviations across the RT and written judgment tasks in Finnish L1 and English L2 (N=66).

In Experiment 2, it had turned out that the learners' and NSs' performance on grammatical and ungrammatical sentences differed (cf Hedgcock 1993). In the present experiment, the learners' judgments and judgment times for correct and incorrect sentences on both Finnish L1 and English L2 were also compared. Figures 37 and 38 show the mean RTs and error scores for both types of sentences.

A series of paired t-tests revealed that for the learners, the difference in the mean RTs of correct and incorrect sentences in the L1 was not significant (t=-1.96, df=65, p=.054), although, as for the NSs of English in Experiment 2, it came close. In the L2, the difference between the correct sentences (mean=2067 msec, sd=436 msec) and the incorrect sentences (mean=2149 msec, sd=404 msec) was statistically extremely significant (t=3.70, df=65, p=.000). The Finns judged both types of L2 sentences significantly more slowly than corresponding L1 sentences. Contrary to the results of Experiment 2, there was a statistically extremely significant difference in the accuracy of the judgments on the correct and incorrect L1 sentences (t=6.30, df=65, p=.000).

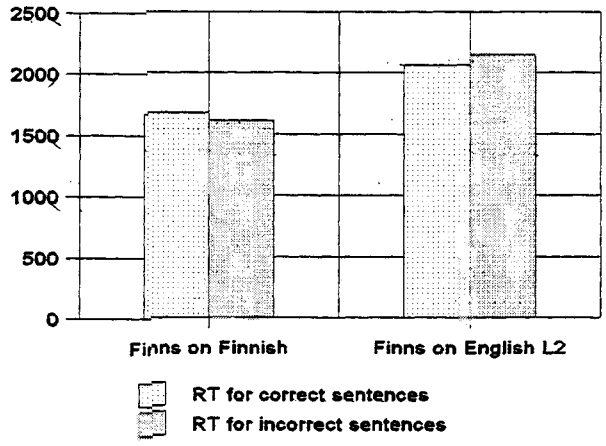


FIGURE 37 The learners' mean RTs on correct and incorrect sentences in Finnish L1 and English L2 (N=66).

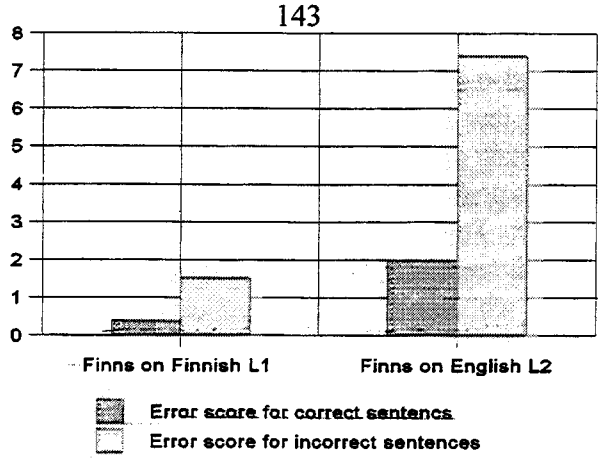
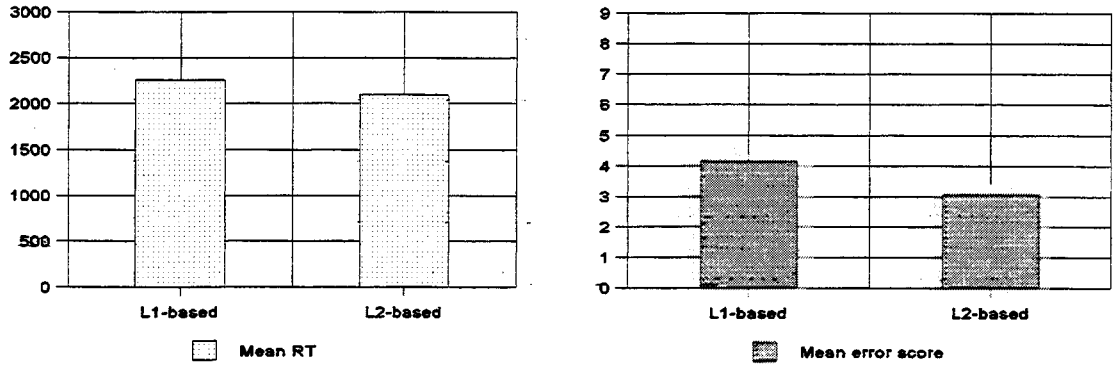


FIGURE 38 The learners' mean error scores on correct and incorrect sentences in Finnish L1 and English L2 (N=66).

(cf. Sutter and Johnson 1990). The difference was much greater for L2, however. Overall, the correct sentences in both L1 and L2 were judged much more accurately than the incorrect ones. It also appeared that the identification and rejection of ungrammatical L2 sentences caused relatively greater problems for the learners.

There were two types of sentences in the experiment: first, there were eight pairs of correct and incorrect sentences that contained errors in the *to* infinitive and 3rd person *-s*. In other words, these were the kind of errors that frequently occur in the speech of English learners, regardless of their background (see eg. Newport and Johnson 1989). Second, there were sentences that were modeled on the L1 of these particular learners. (It has to be borne in mind that errors in the choice of prepositions can be seen to have a system-internal source, as well.) Figures 39a and 39b show the learner performance on these two types of sentences.



FIGURES 39a and 39b. Learner performance on English sentences with L1-based and L2-based errors.

The learners were significantly faster (t=-4.29, df=65, p=.000) in their judgments of the sentences with L2-based errors (mean=2100 msec, sd=420 msec) than sentences with L1-based errors (mean=2261 msec, sd=483 msec). When a t-test was performed on the mean error scores, it turned out that the learners committed significantly fewer errors on L2-based sentences than on L1-based sentences (t=-4.76, df= 65, p= .000).

As is shown in Table 12, the learners were also significantly more accurate in their written English judgments than in the timed English judgment tasks. The learners' responses

to the written judgment task also revealed that they had reacted to the errors the sentences contained. There were a few exceptions in L1 sentences, however: it turned out that the learners' intuitions concerning one infinitive construction in Finnish L1 varied (cf Coppieters 1987). Also, some learners preferred to correct the plural subject *boys* in sentence 113 **The boys loves their little sister.* The correlation coefficient (r=.61) between the accuracy scores of the RT and written judgment tests was statistically significant.

It was also checked whether the order in which the L1 and L2 sentence blocks were presented had any effect on learner performance. A t-test revealed that the order of the languages had no effect on the speed or accuracy of RT judgment data, but it did significantly affect the time spent on both L1 and L2 written judgments, although not their accuracy (see Tables 13 and 14 below).

TABLE 13 The statistical significance of the mean time (in seconds) spent on judging each English L2 sentence in writing and the accuracy of the judgments according to the order of presentation.

Mean Time (in seconds per sentence)	Mean	SD	t	df	p
L2 first (N=33)	13.2	2.99			
			3.37	7.88	.001
L2 second (N=31)	10.3	3.68			
Mean Score	**Mean**	**SD**	**t**	**df**	**p**
L2 first (N=35)	34.4	2.46			
			-.36	58.62	.718
L2 second (N=31)	34.7	2.96			

TABLE 14 The statistical significance of the mean time (in seconds) spent on judging each Finnish L1 sentence in writing and the accuracy of the judgments according to the order of presentation.

Mean Time (in seconds per sentence)	Mean	SD	t	df	p
L1 first (N=30)	12.5	3.34			
			-5.27	43.21	.000
L1 second (N=32)	8.9	1.75			
Mean Score	**Mean**	**SD**	**t**	**df**	**p**
L1 first (N=31)	20.84	.735			
			1.54	61.68	.129
L1 second (N=35)	21.17	1.014			

The results showed that, first, there was a statistically significant difference in the average time that the subjects spent on making grammaticality judgments in writing, depending on the order of presentation. In other words, the subjects seemed to spend significantly more time in judging those sentences they were given first. Second, the order of the tasks had no similar effect on the accuracy the judgments. Consequently, it seems that the subjects became more skilled in making the judgments in writing as the experiment progressed, ie. they got more practice. The order of presentation had no effect on the length of the RT judgments; thus, they appeared to be immune to this type of practice effect. This suggests that there may be a difference in the type of knowledge and processes accessed by these two types of tasks, and that, as Sharwood Smith (1994) suggests, at least this type of grammaticality judgment task may rely on intuitive knowledge or, at least, on processes that are more automatized and less easily accessible to control.

To check for the effects of various background variables, a series of t-tests was carried out. It appeared that whether English was the learners' major or minor subject at the university had no statistically significant effect on the speed and accuracy of RT or written judgments. Whether they were 1st or 2nd year students had no statistically significant effect either. The subjects' gender, however, did have an effect as boys (n=12) were signifantly faster than girls (n=54) in L2 judgments (df=25.41, t=3.79, p=.001) and L1 judgments (df=22.22, t=2.39, p=.026), although the difference for L1 was not as great.

Table 15 shows the average scores of the L2 proficiency test and its subcomponents. Cronbach's alpha for the whole test was .84.

TABLE 15 The subjects' scores on the L2 Proficiency Test and its subcomponents.

L2 Proficiency Test	Mean	SD
Total score of the L2 proficiency test (k=110) (α=.84)	75.1	10.8
Grammar 1 (k=18) (α=.58)	12.5	2.8
Grammar 2 (k=12) (α=.49)	8.6	1.9
Translation (k=10) (α=.09)	7.0	1.2
Reading Comprehension (k=30) (α=.56)	18.8	3.4
Vocabulary (k=40) (α=.76)	28.2	5.3

The proficiency test consisted of three subparts, *Structures*, *Reading Comprehension* and *Vocabulary and Idioms*. For the purposes of the statistical analysis, *Structures* were divided into three parts according to the task type and what sort of information was required to carry them out: *Grammar 1*, *Grammar 2*, and *Translation*. When the internal consistency of the various subparts was calculated, it turned out that *Vocabulary and Idioms* had the highest reliability of all the subparts (α=.76). *Reading Comprehension* and *Grammar 1* were only fair at best. *Translation* had by far the lowest, and *Grammar 2* was

not very reliable either. This was partly due to the low number of items in each subpart. When *Grammar 1*, *Grammar 2*, and *Translation* were combined into *Structures*, the alpha was .70

Table 16 shows the average score of the subjects' answers to Likert-type statements. The range was from 1-5, and the higher the score, the more reward and reliance on their L2 profiency the learners had and the less social anxiety and anxiety by errors they felt. Overall, the learners did not appear to feel particularly anxious about communicating in L2; they also seemed to have a relatively strong reliance on their L2 proficiency. Their L2 learning experience appared to have been quite rewarding as well.

TABLE 16 The average score of the subjects' responses to statements on L2 communicative anxiety (N=61).

	Mean	SD
Reward in English lessons (k=4)	4.0	.68
Reliance On One's Proficiency (k=6)	4.2	.52
Anxiety Caused By Errors (k=6)	3.3	.90
Social Anxiety (k=6)	2.9	.58

10.7.2 Hypotheses 2 - 5

Pearson product-moment correlations were calculated between the variables. The results are shown in Tables 17 - 21. The statistical analysis revealed that:

1. There was a strong statistically significant positive correlation between the English L2 and Finnish L1 judgment times. This was in line with the findings of Experiments 1 and 2.
2. There were no significant correlations between the speed and accuracy of L2 judgments and simple reaction time; ie, those subjects who were slow were not slow because they, in general, had slower reactions. However, the small number of the items in the SRT test may have affected its reliability and validity. On the other hand, the SRT test revealed the usual gender-based difference in RTs: the male subjects (N= 12) were roughly 70 msec faster than the female subjects (N=54). The difference in the mean RTs was not statistically significant ($t=1.74$, $df= 20.65$, $p= .096$).
3. There was no correlation, contrary to the expectations, between the measures of L2 communicative anxiety and the speed and accuracy of the L2 judgments.
4. There was no significant correlation between the speed and accuracy of L2 judgments and the time that the subjects had lived in an English-speaking country.
5. There was a weak but statistically significant negative correlation between the speed of the English L2 judgments and English L2 proficiency test score and a weak but statistically significant positive correlation between the accuracy of English L2 judgments and English L2 proficiency test score.

Analysis of the subjects' performance on the subparts of the proficiency test implied that the accuracy of the L2 judgments (but not the speed) had a stronger statistically significant correlation with the subjects' score on *Grammar 1*. This subpart contained tasks of the following type:

TABLE 9 Correlation Coefficients across L2 and L1 Judgment Tasks. (* = p≤.05, ** = p≤ .01, n.s.= not significant).

	English RTs	Finnish RTs	Simple RTs	English Judgments	Written English Judgments	Finnish Judgments	Written Finnish Judgments	Fluency of Judgments
English RTs	-	.65**	n.s.	n.s.	n.s.	n.s.	n.s.	-.69**
Finnish RTs		-	n.s.	n.s.	n.s.	n.s.	n.s.	-.54**
Simple RTs			-	n.s.	n.s.	n.s.	n.s.	n.s.
English Judgments				-	.61**	n.s.	.32**	-.69**
Written English Judgments					-	n.s.	.35**	.48**
Finnish Judgments						-	.39**	n.s.
Written Finnish Judgments							-	n.s.
Fluency of Judgments								-

TABLE 10 Correlation coeffcienqts across L2 and L1 Judgment tasks and L2 Proficiency Test (* = p≤.05, ** = p≤ .01, n.s.= not significant).

	Total score of the L2 Proficiency Test	Grammar 1	Grammar 2	Translation	Reading Comprehension	Vocabulary
English RTs	-.26*	n.s.	n.s.	n.s.	n.s.	n.s.
Finnish RTs	n.s.	-.30*	n.s.	n.s.	n.s.	n.s.
Simple RTs	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
English Judgments	.29*	.40**	n.s.	n.s.	n.s.	n.s.
Written English Judgments	.32**	.43**	n.s.	n.s.	n.s.	n.s.
Finnish Judgments	n.s.	.31*	n.s.	n.s.	n.s.	n.s.
Written Finnish Judgments	n.s.	.29*	n.s.	n.s.	n.s.	n.s.
Fluency of Judgments	.39**	.44**	n.s.	n.s.	.32*	n.s.

TABLE 11 Correlation coefficients across background variables.

	English Grade in High-School Diploma	Finnish Grade in High-School Diploma	Self-Assesment on Listening	Self-Assesment on Speaking	Self-Assesment on Reading	Self-Assesment on Writing
English RTs	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Finnish RTs	n.s.	n.s.	n.s.	n.s.	-.28*	n.s.
Simple RTs	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
English Judgments	n.s.	.26*	n.s.	n.s.	n.s.	n.s.
Written English Judgments	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Finnish Judgments	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Written Finnish Judgments	n.s.	n.s.	-.26*	n.s.	n.s.	n.s.

Table 12 Correlation coefficients across background variables.

	Age	Years of English Study	Number of Languages Studied	Time of Residence in an English-speaking Country
English RTs	n.s.	n.s.	n.s.	n.s.
Finnish RTs	n.s.	n.s.	n.s.	n.s.
Simple RTs	n.s.	n.s.	n.s.	n.s.
English Judgments	n.s.	n.s.	n.s.	n.s.
Written English Judgments	-.25*	n.s.	n.s.	n.s.
Finnish Judgments	-.36**	n.s.	n.s.	n.s.
Written Finnish Judgments	-.29*	n.s.	n.s.	n.s.

TABLE 3 Correlation Coefficients for the L2 sentences containing intralingual errors and errors beased on L1.(* = p≤.05, ** = p≤ .01, n.s.= not significant).

	English RTs on incorrect sentences (L2)	English RTs on incorrect sentences (L1)	English Judgments on incorrect sentences (L2)	English Judgments on incorrect sentences (L1)
English RTs	.94**	.90**	n.s.	n.s.
Finnish RTs	.65**	.59**	n.s.	n.s.
SRT	n.s.	n.s.	n.s.	n.s.
English Judgments	n.s.	n.s.	.73**	.72**
Written English Judgments	.n.s.	n.s.	.58**	.38**
Finnish Judgments	n.s.	n.s.	n.s.	n.s.
Written Finnish Judgments	n.s.	n.s.	.26*	.33**
Fluency of Judgments	-.69**	-.62	.52**	.30*
Total score of the L2 Proficiency Test	-.27*	n.s.	.41*	n.s
Grammar 1	n.s	-.26*	.49**	n.s.
Grammar 2	n.s	n.s.	n.s.	n.s.
Translation	n.s	n.s.	n.s.	n.s.
Reading Comprehension	-26*	n.s.	n.s.	n.s.
Vocabulary	n.s.	n.s.	.33**	n.s.
Social Anxiety	n.s.	n.s.	.27*	n.s.

Example 5 A tax declaration form is _____ for a small child to understand.
(a) too difficult a thing
(b) a too difficult thing
(c) too difficult thing

This type of task was somewhat similar to the grammaticality judgment task in that in both tasks the subjects had to focus on a particular linguistic form within a sentence. The difference is, however, that this type of task is productive in a very limited sense, ie. the subject has to choose which of the options she would rather use in the sentence in question. Both tasks seem to involve search in memory through grammatical knowledge. The correlation may be explained by the difference among learners in how good they were in searches like these. It is remotely possible that the correlation, especially since it was rather weak, could have been at least partly a product of task type (see eg. Hulstijn 1989). However, the analysis of individual differences in learner performance seems to exclude this possibility.

6. There was a significant correlation between the accuracy of the RT and written judgments. On the whole, however, the subjects were significantly more accurate in their written judgments. They were better able to access knowledge 'that' and very likely explicit formal metalinguistic knowledge during the written test. The difference in results may indicate that during the RT task, under testing conditions, the subjects seemed to rely on a different source of knowledge in order to be able to judge the sentences as quickly as possible.

10.7.3 L2-based errors vs. L1-based errors

The results revealed interesting differences in the RT and accuracy data from the two types of incorrect sentences included in the experiment. Table 13 shows all the correlation coefficients for these two types of sentences. When the Pearson product-moment correlations (shown in Table 22) were calculated by using learner performance on these two sets of sentences as data, it appeared that the speed and accuracy of the judgments of L2-based errors correlated more significantly with various L2 proficiency scores than did the RTs for the whole set of sentences. Thus, the mean RT for these sentences was weakly related to the total score on the L2 proficiency test (r=-.27*) and also to the reading comprehension score (r=-.26*). Moreover, the accuracy of the subjects was positively correlated with their score on written English L2 judgments (r=.58**), the L2 proficiency test as a whole (r=.41**), *Grammar 1* (r=.49**), and *Vocabulary and Idioms* (r=.33**). Furthermore, it was also related to the subjects' score on written Finnish L1 judgments (r=.26*).

The expected correlation with the affective variables of L2 communicative anxiety also materialized: SOCANX was weakly, yet significantly, related to the accuracy on these sentences. In other words, there appeared a slight tendency for the subjects who reported themselves as more confident in social situations to be more accurate in detecting this kind of error.

10.7.4 Fluency

A new sum variable called FLUENCY was also formed to better capture the combined effect of the speed and accuracy of the subjects' L2 judgments. This was accomplished by

first calculating the z scores of the mean judgments and judgment times, then reversing the judgment time scores by multiplying the original z score by -1 and adding these two z scores up. In other words, the faster and more accurately the subject had performed, the higher the fluency score was. The correlation coefficients linked to FLUENCY are shown in Tables 17 -20.

When the correlation of this new variable FLUENCY with other variables was calculated, it was revealed that:

1. There was a relatively strong, statistically significant negative correlation between FLUENCY and the speed of the subjects' L1 judgments. In other words, the higher the subjects' fluency, the faster they were in judging L1 sentences (not unexpected considering the strong correlation beteen L1 and L2 speeds).
2. FLUENCY had a significant correlation with the subjects' accuracy in written L2 judgments (again, not unexpected considering the correlation between the accuracy of timed and written L2 judgments).
3. FLUENCY had a significant correlation not just with the subjects' total score on the L2 proficiency test and their score on *Grammar 1* but also with their score on *Reading Comprehension*.

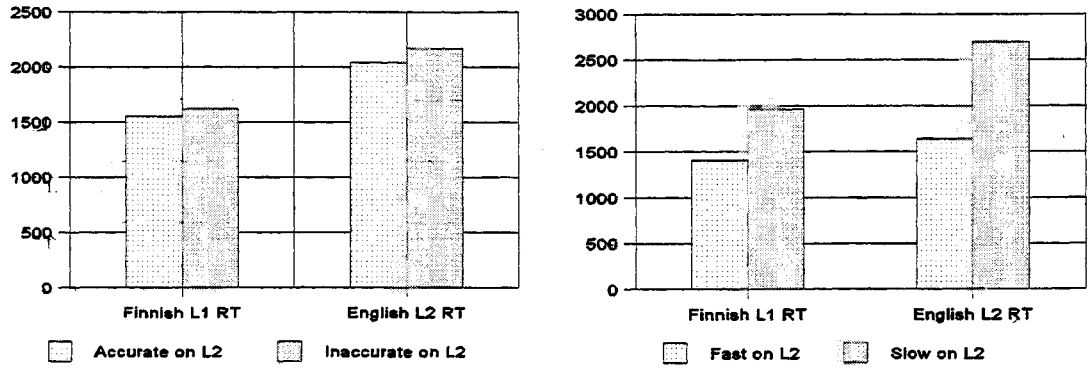
A multiple regression analysis revealed that the speed with which the subjects made their L1 judgments accounted for 28% for the variance in FLUENCY. The speed of L1 judgments together with their score on *Grammar 1* accounted for 37% for the variance in FLUENCY. The regression coefficient was statistically significant in both cases.

These findings seem to suggest that there is a language-independent element in this type of task that may be inherent in the task requirements. After all, the successful execution of this task in both L1 and L2 requires focused attention on linguistic form during an on-line task. The significant although rather weak correlation with the L2 proficiency test score, especially that section of it which required the subjects to search their memory for a correct grammatical structure to fill in the L2 sentence, and in the case of FLUENCY, with *Reading Comprehension*, seems to suggest that if this type of judgment task has a connection with general L2 proficiency, it is through the nature and degree of automatization of linguistic and metalinguistic knowledge in language performance.

10.7.5 Individual learner performance: 'Slow' and 'Fast' judges

To help to characterize the subjects' performance in this type of task, they were grouped into upper and lower 25% according to the speed and accuracy of their L2 judgments. The results are shown in Figures 40 - 42. A comparison of the means showed that:

First, the 'Accurate' judges were better than the 'Inaccurate' not only in the RT judgments but also in the written judgment tasks, as well as in most parts of the language proficiency test; at the same time, they were not significantly slower than the 'Inaccurate' ones. Secondly, 'Fast' judges were no less accurate than 'Slow' judges in the timed judgment tasks nor in the written judgment tasks. In other words, there appeared to be no trade-off between speed and accuracy for these learners; rather, their perfromance was similar to fluent, skilled performenca. Also, the 'Fast' judges were almost significantly better in the reading comprehension part of the language proficiency test. This seems to suggest that there could be a connection between speed and accuracy, ie. fluent, skilled performance in the L2 judgment tasks and L2 proficiency in general.



FIGURES 40a and 40b The performance of 'Accurate' (N=5) and 'Inaccurate' (N=5) and 'Fast' (N=5) and 'Slow' (N=5) learners on the judgment task.

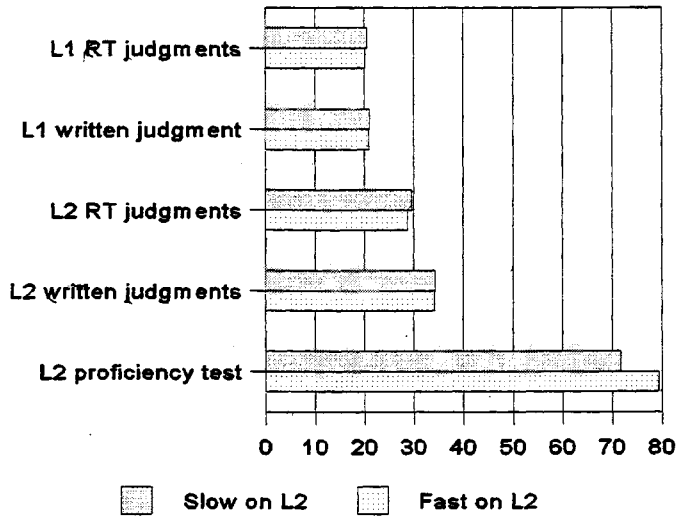


FIGURE 41 The performance of 'Fast' and 'Slow' judges on the subparts of L2 proficiency test.

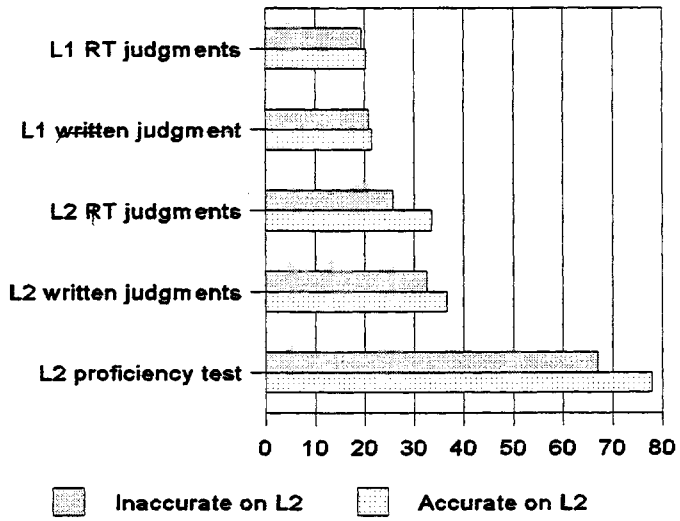


FIGURE 42 The performance of 'Accurate' and 'Inaccurate' judges on the subparts of L2 proficiency test.

10.7.6 Individual learner performance: cluster analysis

To obtain a better view of individual differences, the results were submitted to a cluster analysis. According to Skehan (1986, 1989), cluster analysis can be a useful tool for analyzing individual differences. First, the variables of the speed and accuracy of the L2 judgments, L2 proficiency score and anxiety about errors were chosen for the basis of

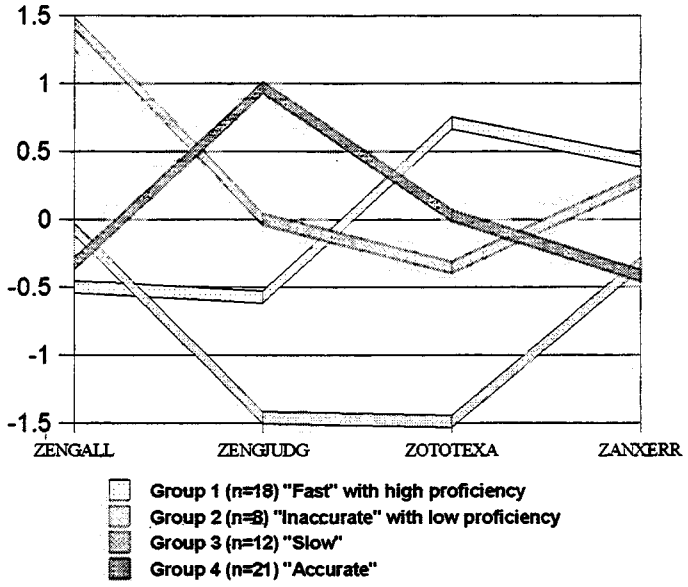


FIGURE 43 Learner subgroups and their profiles on variables used in the cluster analysis

cluster analysis and transformed into z-scores. Various cluster solutions were obtained by using Ward's method. A visual inspection of the resulting dendrograms suggested a four or five cluster solution. Of these, the four cluster solution was chosen to be presented here. On the basis of the solution, the following subject groups were formed: Group 1 (N=18), Group 2 (N=8), Group 3 (N=12) and Group 4 (N=21). A oneway ANOVA with a post hoc Scheffé procedure was carried out on the variables included in the cluster solution to detect the significant differences between the new subject groups and to better describe the subject clusters.

The ANOVA revealed that there was a significant difference between the groups as to the mean RT of the English L2 judgments (F=20.33, p=.000). A post hoc Scheffé procedure showed that Group 3 was significantly slower than the other groups in their L2 judgments. There was a significant difference in the accuracy of the L2 judgments (F=45.60, p=.000) as well. A post hoc Scheffé procedure showed that Group 2 was significantly less accurate than the other groups and Group 4 significantly more acccurate in their L2 judgments. The ANOVA also showed that there was a significant difference among the groups as to the total score of the L2 proficiency test (F= 19.73, p=.000). According to the results of a post hoc Scheffé test, Group 2 scored significantly lower and Group 1 significantly higher than the other groups. Finally, the ANOVA showed that there was a significant difference, although not nearly as striking, among the groups as to the anxiety about errors (F=3.25, p=.0287). However, no group differed from each other to

a significant degree. The groups can thus be characterized as follows:

Group 1: higher proficiency L2 learners (hence labeled as 'Good')
Group 2: less accurate and lower proficiency L2 learners (hence labeled as 'Poor' and 'Inaccurate')
Group 3: subjects who were slow (hence the label 'Slow')
Group 4: language learners who were more accurate than others in the L2 judgment task (labeled as 'Accurate').

It became immediately obvious that there was a difference between those subjects who scored high in the L2 proficiency test and those whore were classified as 'Accurate' in the RT judgment task. Although there was a correlation between the speed and accuracy of the L2 grammaticality judgments and some parts of the L2 proficiency test, the ones with higher proficiency were not necessarily the ones who were the most accurate in this type of task. This does not mean that they were inaccurate, rather that they were not as good as Group 4. However, this seems to indicate that the grammaticality judgment tasks might, indeed, measure a specific type of grammatical knowledge and skill that might be only part of a more general measure of L2 proficiency.

Furthermore, there is no such group as 'Fast', rather there was simply a group of learners that was slower than others. This is in line with the results from Experiments 1 and 2 that showed that there was a much greater consistency among the slow performers than the fast ones (the same subjects turned out to be consistently slower, while there was some variation as to who was ranked among the fastest 25% in each subject group).

To validate the solution, a oneway ANOVA with a post hoc Scheffé procedure was carried out on four other variables. Figure 44 shows the profiles of the learner subgroups on the variables used to validate the cluster solution. The accuracy of the L1 judgments was so high that it was not expected to produce any differences between the groups. Instead, the accuracy of the subjects' L2 judgments (ZNEWEJUD) in the written task was included in the procedure. The groups appeared to remain relatively distinct in their characteristics:

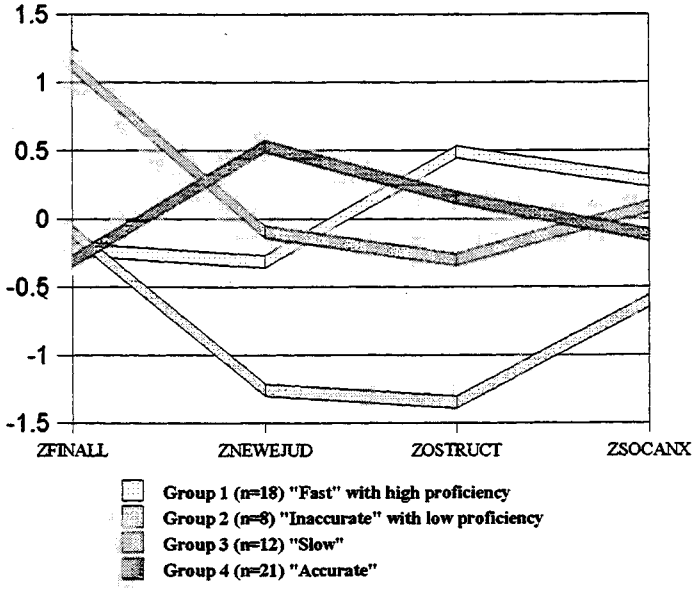


FIGURE 44 Learner subgroups and their profiles on variables used to validate the cluster solution.

The ANOVA revealed that there was a significant difference among the groups as to the speed of the L1 judgments (F=7.72, p=.000), with Group 3 being significantly slower than others even in judgments involving their Finnish L1. The mean scores of the L2 written judgments also differed significantly form each other (F=10.55, p=.000). A post hoc Scheffé procedure showed that Group 2 was significantly less accurate than Groups 3 and 4, while Group 4 was significantly better than Groups 1 and 2. There was a significant difference in the *Structures* subpart (consisted of *Grammar 1*, *Grammar 2* and *Translation*) of the L2 proficiency test (F=10.31, p=.000), with Group 2 scoring significantly lower than Groups 1 and 4 (but not Group 3). No signicant differences were apparent among the groups as to social anxiety.

The ANOVA was carried out to explore the nature of the differences among the groups in all the subparts of the L2 proficiency test as well. *Translation* turned out not to differentiate among the subject groups. *Grammar 1* produced a significant difference, however (F=14.04, p=.000). A post hoc Scheffé procedure showed that Group 2 scored significantly lower than the other groups and that Group 1 significantly higher than both

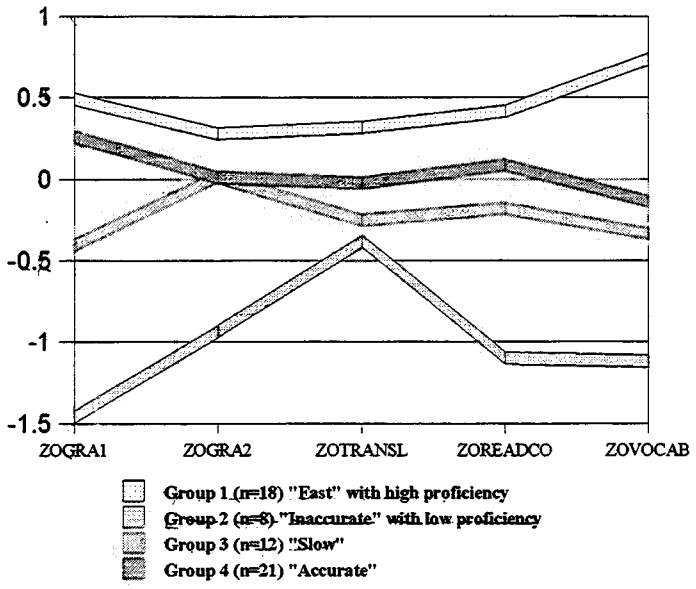


FIGURE 45 Learner performance on the subparts of the L2 proficiency test.

Groups 2 and 3. *Grammar 2* also produced a significant effect although not to the same degree (F=3.15, p=.032). This time Group 1 was significantly better than Group 2. Reading Comprehension produced a significant effect as well (F=5.61, p=.002), with Group 2 having a lower performance than Groups 1 and 4. The ANOVA revealed that the difference in the *Vocabulary* scores was significant (F=11.43, p=.000). There was a very clear advantage in *Vocabulary* for Group 1: it scored significantly better than all the other groups. Group 2, again, scored lower than both Groups 1 and 4.

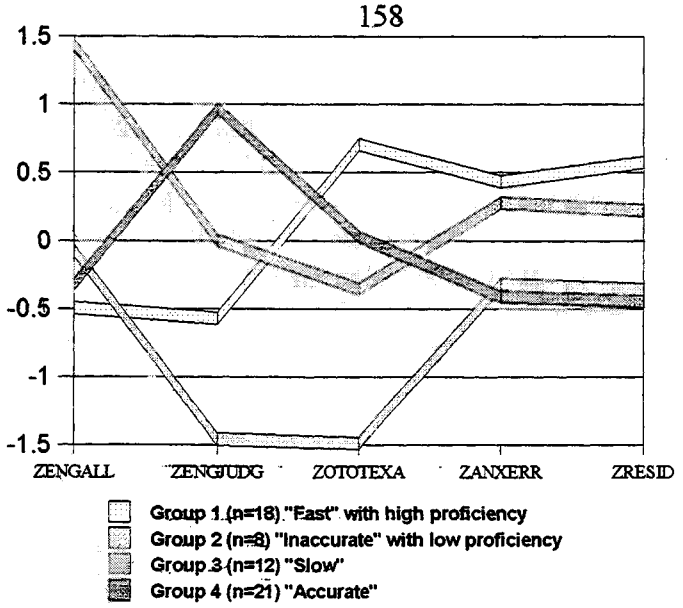


FIGURE 46 The relationship of 'Time of residence' to the variables used in the cluster solution.

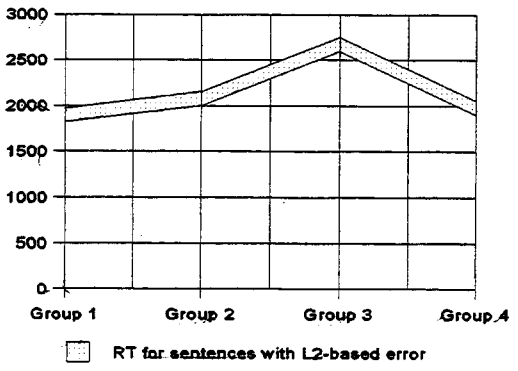
Background variables such age, the number of languages studied, or the years of English study produced no significant effects among the groups. When the learners' group membership was checked against the background variables, certain interesting differences were revealed among the groups. It turned out that Group 1 (N=18) had 12 subjects (67%) who had, unlike most of the other groups, no other modern language as their minor subject. Instead, they had subjects such as political science, computer science, ethnology, ancient Greek, or literature, or no minor subject at all. All except one subject in this group were also English majors. Group 2 (N=8), on the other hand, turned out to contain six students (75%) who had some other language (such as German, Swedish, French and Russian) as their major language. Thus, although a paired t-test showed that for the whole group the difference between English majors and minors was not significant as to the learner performance on the judgment task, it turned out to play a role, very likely through a lower English proficiency these learners had. Group 3 (N=12) and Group 4 (N=21), however, show no distinguishing characteristics. 17% of Group 3 and 19% of Group 4 were English minors. 58% of Group 3 and 67% of Group had another modern language as their other subject. Table 21 shows the learners' major and minor subjects and the time of residence in an English-speaking country. When a oneway ANOVA was carried out as to the background variable 'Time of residence in an English-speaking country', however, an interesting difference was revealed: it turned out that there was a significant difference between the groups ($F=4.5$, $p=.007$), with Group 1 having stayed significantly longer than Group 4 in an English speaking country. (Again, it is to be noted that the Pearson product-moment correlation coefficients did not indicate that the time of residence had a significant relationship with any of the variables.) Moreover, 76% of learners in Group 4 had spent only one month or less in an English-speaking country. As a matter of fact, as many as 48% had spent no time at all in an English-speaking country within the last ten years. In Group 1, 22% of the learners had spent one month or less, and only 11% no time at all. Figure 46 shows the relationship of the learners' speed and accuracy of L2 judgments, L2 proficiency test score, anxiety about errors and time of residence in an English-speaking country. It was

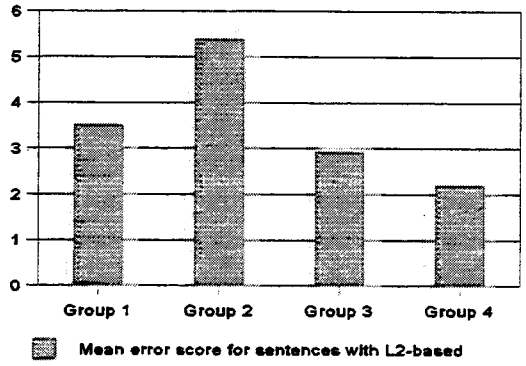
TABLE 21 Learner subgroups and their relationhip to the background variables of major and minor subjects and the times of residence in an English-speaking country.

		English majors	Modern language minors	Average time of residence
Group 1: (N=18)	'Good'	94%	33%	6.4 months
Group 2: (N=8)	'Inaccurate' and 'Poor'	25%	88%	2.3 months
Group 3 (N=12)	'Slow'	83%	58%	4.3 months
Group 4 (N=21)	'Accurate'	86%	57%	2.2 months

an important finding since this was the first indication that there might be a difference in the nature of L2 proficiency between these groups. The finding takes on even more meaning when it is combined with the finding that Group 1 contained a much larger proportion of subjects who studied no other modern languages, and that English was their major subject. This is another indication that the whole language learning history of these subjects might be different.

To further characterize these four learner groups, their mean RTs and error scores were calculated on the English L2 sentences that contained L2-based errors, since they





FIGURES 47a and 47b The performance of the four learner groups on English L2 sentences with system-internal errors.

appeared to be most clearly related to L2 proficiency and social anxiety variables. Figures 47a and 47b show the results of the descriptive statistics.

The ANOVA was carried out to see how their judgments times differed on the English L2 sentences that contained an L2-based error. Table 22 shows the source of variance for these sentences.

Source	DF	Sum of squares	Mean squares	F	p
Between groups	3	4910531.473	1636843.824	15.62	.000
Within group	55	5763278.322	104786.8786		
Total	58	10673809.80			

TABLE 22 Source of variance, F-value and probability for the mean RT on English sentences containing an L2-based error (k=8) for Group 1 'Good' (N=18), Group 2 'Inaccurate and Poor' (N=8), Group 3 'Slow' and Group 4 'Accurate' learners of English.

A post hoc Scheffé procedure revealed that Group 3 was significantly slower than the other groups. The ANOVA was also carried out to see how their judgments on the English L2 sentences that contained an L2-based error differed. Table 23 below shows the source of variance for these sentences.

Source	DF	Sum of squares	Mean squares	F	p
Between groups	3	61.7753	20.5918	22.414	.000
Within group	55	50.5298	.9187		
Total	58	112.3051			

TABLE 23 Source of variance, F-value and probability for the mean error scores on English sentences containing an L2-based error (k=8) for Group 1 'Good' (N=18), Group 2 'Inaccurate and Poor' (N=8), Group 3 'Slow' and Group 4 'Accurate' learners of English.

A post hoc Scheffé procedure revealed that Group 2 was significantly less accurate than the other three groups; in addition, Group 1 was significantly less accurate than Group 4 on these sentences. This finding is very interesting because it raises the question why the learners who had spend significantly more time in an English-speaking country, and scored higher expecially in the vocabulary section of the L2 proficiency test, did not perform as well on these sentences as the learners who had a more typical foreign language learning background in that they had not spent as much time as Group 1 in a naturalistic language learning environment. This can be a further indication that this type of judgment task requires of a specific metalinguistic skill.

10.8. Conclusion

In the beginning, a series of hypotheses were made as to the relationship between the speed and accuracy of the RT grammaticality judgments and L2 proficiency, on the one hand, and L2 communicative anxiety, on the other. The results showed that the reaction time task with its two dimensions of speed and accuracy appeared to reflect some aspects of language proficiency, especially a certain type of grammatical knowledge.

The statistical analysis of the results indicated, among other things, that the students who scored higher in the language proficiency test tended to be faster and more accurate in their judgments of grammatical acceptability. When the students' performance was analyzed in more detail, it appeared that the better their performance in a grammar section of the language proficiency exam was, the more accurate they were on their judgments of

both English and Finnish sentences. No such correlations were found with their performance in the reading comprehension or vocabulary section. However, when the same correlation analyses were conducted with only certain types of erroneous sentences, some links appeared between the variables. This suggests that this type of linguistic task is related to more general measures of language proficiency. The results also appear to indicate that the link between the performance on the RT grammaticality judgment task and the performance on the L2 proficiency test is through the notions of automaticity and the application of automatic procedural skills. This was further supported by systematic individual differences across the RT latencies: the students who were fast in their judgments of L1 tended to be fast in L2 as well.

Differences in the application of such skills might also explain the intriguing differences found among the learners: the cluster analysis revealed that the learners' task performance varied significantly according to background variables such as the time spent in an English-speaking country and the number of languages studied. This was important from the methodological point of view as well: the Pearson product-moment correlations had failed to show significant correlations between the task performance and these two background variables. The factor underlying the two background variables can only be assumed. It does appear, though, that it involves the recent language learning history of the learners.

There also appeared to be a metalinguistic or general cognitive skill element involved. First, the RTs for both languages included in the experiment were highly correlated. Second, as mentioned above, the learner group with a clearly formal language learning background outperformed the learner group with a more naturalistic language learning history. However, unlike the linguistic and metalinguistic knowledge and skills applied in the written judgment tasks, the skill element involved in the RT judgment tasks appeared to be immune to practice effect since the judgments remained unaffected by the order of presentation of the sentences. It may have been that the learner group with a more formal language learning background were inclined to focus their attention to form during language performance.

11 Discussion and conclusion

In this thesis, an attempt was made to investigate factors and processes involved in a RT grammaticality judgment task. A psycholinguistic model of second language performance was used to provide a framework within which the relationship between cognitive and affective variables, L2 knowledge, and the speed and accuracy of RT judgments of grammatical acceptability was examined. Within a such framework, language proficiency can be defined as an ability to use the mechanisms involved in language performance in a way which results in fluent skilled language behavior. RT judgments can be thought to provide important information about the functioning of such mechanisms.

To sum up the results of the three experiments, it was found that on an average, the NSs were almost always significantly faster and more accurate in their RT judgments than the language learners. In Experiments 2 and 3, when the Finns were asked to judge Finnish L1 sentences, it turned out that their reactions were faster and more accurate for the L1 sentences than for the L2 sentences. In Experiments 2 and 3, the results also indicated that the learners' responses to ungrammatical L2 sentences showed a great deal more variation than their responses to grammatical L2 sentences; in this respect, learner performance differed from NS performance, as the NSs showed less variance between the two categories of sentences. The RT judgments also showed the effects of transfer. In most cases this appeared best in learners' reactions to ungrammatical sentences. The L1s of the learners appeared to influence their performance on English L2; the Finns' responses to German L3 sentences also suggested their willingness to transfer structures from the stronger English L2 into the weaker language. Certain relatively consistent tendencies also emerged among the individual learners. Thus, it appeared that the same subjects were among the slowest performers across different languages.

Of the cognitive and affective variables included in the study, only some aspects of L2 communicative anxiety appeared to be related to the RT judgments. In Experiment 3, the subjects' L2 proficiency was also measured. Weak yet statistically significant correlations appeared between the speed and accuracy of RT judgments and one of the grammar sections of the L2 proficiency test, in particular. This tendency became clearer when only the responses to a certain type of ungrammatical L2 sentences were included in the study. The cluster analysis revealed a number of intriguing differences among the learners as well.

In all experiments, the subjects' mean RTs correlated with each other across the languages. In addition, a group of language learners emerged in Experiment 3 tha appeared to be more accurate in the RT judgments than other type of learners. When the background variables were examined, it appeared that this group of learners had been exposed to English primarily in formal language learning settings. Findings such as these suggest that there is a metalinguistic skill element involved that has most likely to do with focus on form.

11.1 RT judgments and L2 skilled performance

It was assumed in the beginning that native speakers' performance on this type of task when it involved their L1 was a reflection of their higher language proficiency. In other words, native speakers would be faster and more accurate in the RT judgments of grammatical acceptability. A distinction was made between automaticity, on the one hand, and skilled performance, on the other hand, the former being regarded primarily as a function of speed, while the latter was considered to be essential for both the speed and accuracy of judgments.

In Experiment 1, there were two groups of native speakers, NSs of English and NSs of German, whose performance on their L1s was compared with that of Finnish language learners. For both English and German sentences, the NSs were significantly faster and more accurate, but the NSs were not perfect either. Both of these findings are hardly unexpected. However, this is an indication that the unreliability and variability discovered in the NS judgments of grammaticality is only relative (Coppieters 1987, Nagata 1987, 1988, 1989a, 1989b, 1989c); a point that perhaps needs to be made more clearly.

Although the German proficiency level of the Finnish language learners was not explicitly measured in Experiment 1, the *post hoc* comparisons between the students of German and non-students of German revealed a significant difference in the RTs on the judgments of both the German sentences and the English/German translation pairs. The difference was quite remarkable, as the performance of the two *post hoc* groups in the other language blocks was almost indentical. Thus, the longer RTs of the non-students of German could not be explained by assuming, for example, that they were generally slower in this type of task, since there were parctically no differences between their RTs for the other languages. The students of German were also significantly more accurate than the non-students in judging German sentences. The difference could only be explained by the students' different proficiency levels in German. The underlying causes for the observed difference could be traced to several sources. One such explanation could be a difference in the reading skill and one of its most basic element, visual word recognition. The students of German, by simply having a greater contact and familiarity with the language, have most likely developed a higher level of automaticity in reading German (see eg. Favreaux and Segalowitz 1983, Segalowitz 1986). There must have also been an overall difference between the learners' declarative and procedural knowledge of the German language. In terms of Anderson's ACT* model of skill learning, the non-students' skills had not been proceduralized to the extent that the German students were.

In Experiment 2, there was an almost significant difference in the RTs and a significant difference in the accuracy between the NSs of English and the learners. The speed with which the Finnish learners made their judgments was in this experiment closer to that of the NSs; in this respect, the behavior of the Finnish students of English in Experiment 2 bore a closer resemblance to that of the German students of English in Experiment 1. The NS performance again had the characteristics of automatic skilled performance; the learners, while being almost as fast as the NSs in their judgments of the control sentences in particular, were less accurate. The results could be taken as an indication of the influence of L1-based procedural and declarative knowledge on the processing of these sentences. This seems to be the case especially for structures for which the deviant sentence was judged as grammatical and the nondeviant sentence as ungrammatical.

In Experiment 3, the language learners' performance was compared, first, with their English L2 proficiency level. It appeared that there was a weak, yet significant negative

correlation between the RT and the total score of the proficiency test. In other words, there was a slight tendency for those learners who scored higher in the proficiency test to be faster in their judgments. Although no overall correlation appeared between the accuracy of the judgments and the total score of the proficiency test, the scores for a subsection of the test, focusing on grammatical structures, turned out to be positively related to the accuracy of judgments. This tendency was even clearer for the learners' speed and accuracy in judging deviant sentences of a certain type; the correlation coefficients were higher and more significant between the various subparts of the proficiency test and the RT and judgments for these sentences. This seemed to indicate that it was the aspect of the RT judgment task that required the learners' to detect deviancy that was related to L2 proficiency, or at least this particular measure of L2 proficiency. Indirect support for this comes from Hedgcock (1993), who points out how language learners tend to judge well-formed sentences more consistently and accurately than ill-formed ones. The relatively high level of L2 proficiency may play a role here. It stands to reason to assume that the more advanced the learners are, the faster and more accurately they will be able to judge grammatical sentences, in which case the learners' responses to ungrammatical sentences may be more indicative of their interlanguage knowledge. It may be that for the more advanced learners, the differences in language proficiency may become apparent by examining their responses to ungrammatical sentences.

Yet, the ability to detect deviancy fast and more accurately seems to be an aspect of the RT judgment task that is also influenced by the metalinguistic skills of the subjects. When a cluster analysis was performed on the results of the Finnish learners in Experiment 3, a four-cluster solution emerged that revealed interesting systematic differences between the learners. One group of learners (Group 1) scored higher than the others on the L2 proficiency test; yet, they were not the fastest and most accurate group of learners in the RT judgment task (although they were better than two other groups). The group (Group 4) which was fastest and most accurate primarily consisted of learners who had spent significantly less time in an English-speaking country than the group with the higher proficiency level. Many of the learners in this group had spent either no time at all or less than one month in an English-speaking country. Another difference that emerged was the number of the languages studied: most of the learners with a higher level of proficiency studied only one modern language, English, at the university. The results were also interesting from the methodological point of view: no significant Pearson product-moment correlations had appeared between the RT judgment and proficiency variables and background variables of the time of residence in an English-speaking country or the number of the languages studied. Yet, after a cluster analysis, these two background factors emerged as important intervening variables.

One explanation for this difference in performance could be a greater focus on form by the learners in Group 4, brought on by the formal language-learning setting, for example (cf. Hedgcock 1993: 3). The learners in Group 1, on the other hand, may have carried out the task by using a more meaning-based approach and perhaps paid less attention to form. The fact that the majority of Group 1 studied English as the only foreign language, while having subjects such as psychology and political science as their minors, may also indicate the existence of underlying differences in the whole language learning history and motivation for these two subject groups (cf. Birdsong 1989b).

There was a significant positive correlation between the RTs for the sentences in different languages. This provides further support for the assumption that there is a language-independent metalinguistic skill element involved. However, as it turned out in

Experiment 3, although the subject group that was labeled as 'Slow' on the basis of their L2 performance was slow on the L1 judgments as well, they were not as accurate as some other learners. Hence, it appears that the longer time such 'slow' subjects spent on these tasks did not result in a higher level of accuracy. It also appears that these subjects may have a tendency to monitor or search for explicit knowledge from long-term memory during the test performance, a tendency which also appeared in L1 (or L3 and L4) performance. There were individual differences across the experiments, though, as some subjects were remarkably fast in L1 but not in L2 or L3.

Experiment 1 also offered an opportunity to compare the performance of two different NS groups: German teachers and German students. The results showed that the German students were significantly faster that the German teachers in judging the German sentences, but there was no significant difference in their accuracy. As to the English/German translation pairs, the students had again significantly lower RTs but this time they were significantly less accurate than the teachers. In the latter case, but not in the first, there may have been a trade-off between speed and accuracy. The performance of the German students seemed to have the characteristics of automatic skilled performance in the first type of judgment task, but in the second type of judgment task, their accuracy fell off below that of the German teachers. The task may have demanded a specific type of metalinguistic skill that the teachers, perhaps because of their background, had mastered better.

11.2 Factors affecting RT judgments

Lehtonen (1984) and Bley-Vroman and Masterson (1989) have discussed the research techniques in conducting reaction time research in combination with acceptability or grammaticality judgments (see also Hulstijn 1997). The topic also comes up in Birdsong (1989a), Ellis (1991), Hedgcock (1993), Cowan and Hatasa (1994) and Gass (1994). In Chapter 3, a group of variables that may affect a grammaticality judgment task were presented: first, the nature of the linguistic stimuli; second, the cognitive and affective variables associated with the judge himself or herself; and third, the test procedure and situation. In the following, the results of the three experiments will be related to these factors.

1. The nature of linguistic input, ie. the item to be judged. It appeared that grammatical (or well-formed, non-deviant, or correct) and ungrammatical (ill-formed, deviant, incorrect, or erroneous) sentences were judged with different accuracy and speed. Subjects had a low rate of accuracy on ungrammatical sentences. Most of the ungrammatical sentences were based on errors typically produced by the Finnish L1 learners of English. That the Finnish learners showed great variation in the judgments of such sentences is an indication that either their interlanguage knowledge or processes may not have stabilized. In terms of declarative and procedural knowledge, it is difficult to attribute the subjects' performance to either one of them unless it can be shown that the subject (at least on this occasion) systematically rejects the grammatical sentence and accepts the L1-based ungrammatical sentence. If this is the case, it can be said with more confidence that the response has its source in the declarative knowledge base. There were a few prepositional structures in Experiment 2 (eg. 102 **Amy had a golden ring in her finger* vs. 108 *You have a new ring on your finger*) for which some subjects exhibited this type of behavior. A much larger

proportion of judgments revealed a tendency by the learners to rely on what could be called procedural knowledge developed during the course of formal language learning. The subjects may have assumed the equivalency of certain L1 and L2 forms (eg. *-ssa = in* and vice versa) and used this knowledge to build procedures at some point of the language learning process. In some cases these procedures appear to have become quite automatic as indicated by fast acceptance of L1-based structures in L2.

It also appeared that the nature of errors played a role as the judgments involving prepositional constructions based on L1 or certain system-internal errors as in the infinitive structure turned out to be extremely variable. Some of the structures based on Finnish L1 (as appeared in Experiment 1), however, turned out to be easy to recognize for the Finns, but not so easy for the Germans (sentence 106 **He left it to Jill's room* is a good example). The effect of metalinguistic awareness in terms of perceived distance between languages (Kellerman 1977) or well-known garden-path sentences (cf. Tomasello and Herron 1989), brought on by formal language instruction may have affected learner performance.

2. The judge himself or herself. In Experiment 2 and, to a lesser degree, in Experiment 3, the subjects' performance in the RT judgment task was related to various affective factors such as field-dependence/ independence, verbal and nonverbal thinking styles, and language anxiety. Of these, field-dependence/independence and thinking style turned out to have no relation to the speed and accuracy of L2 judgments in Experiment 2. In Experiment 3, the various factors of language anxiety had no relation to the overall speed and accuracy of the RT judgments for the whole set of sentences; it turned out, however, that the L2 sentences containing a system-internal error were judged slightly more accurately by the subjects who felt less social anxiety.

3. The test procedure and situation. In Experiments 1 and 2, the ungrammatical sentences were always presented first. Although the experiments were conducted under time pressure (for example, the sentences were visible only for a very short time) and great care was taken to ensure that a sufficient number of other sentences lay between similar grammatical and ungrammatical structures, the fact that the subjects had seen the ungrammatical sentence first may have affected the speed and accuracy with which they judged the sentences. In Experiment 3, the sentences were presented in random order, yet a similar tendency appeared as L2 grammatical sentences were judged with greater speed and accuracy.

11.3 RT judgments and the use of L2 knowledge

Intially, it was assumed that the RT judgments were a reflection of the subjects' use of automatic or controlled processes: the faster the reaction, the more automatic the response. No firm assumptions were made as to the nature of knowledge: it was simply assumed that longer RTs would be an indication of search in long-term memory. Indirectly, however, by allowing the subjects only 2000-2100 msec to view each sentence, it was hoped that their decisions would be more intuitive and perhaps in that sense based on implicit knowledge.

The subjects' performance can be interpreted by means of Anderson's ACT* theory of skill learning. In Anderson's theory, the declarative stage of skill learning is roughly equivalent, but not identical, to controlled processing and procedural stage to automatic processing in the sense Shiffrin and Schneider (1977) use these terms. These two types are

knowledge are also indirectly related to implicit and explicit knowledge. Many foreign-language learners, who often receive formal language instruction, use explicit grammar rules as the starting point for the establishment of automatic routines (cf. Hulstijn 1990: 38). It also appeared that automatic routines may coexist with explicit knowledge of rules. During the course of foreign language learning, after having established the relevant automatic routines, learners may have forgotten the rule itself. Yet, they may also recall it. As Hulstijn (1990: 35) points out, it is possible for both procedural knowledge and declarative knowledge to exist side by side. However, the skilled performance is still governed by procedural knowledge (Anderson 1980: 226).

Schmidt (1992: 364- 365) suggests that Anderson's model relates to fluency (defined by Schmidt as automatic procedural skill, ie., how well procedural knowledge is actually used in real-time situations) in various ways: proceduralization contributes to the speed of processing, since it decreases the load on working memory while freeing resources for other tasks that may be performed simultaneously. Similarly, Towell, Hawkins and Bazergui (1996) also argue that the proceduralization of linguistic knowledge is the most important factor in the development of fluency at least in advanced second language learners. Moreover, it has to be assumed that what applies to language performance applies to the metalinguistic aspect of that performance, as well; what applies to the learning of language skills applies to the learning of metalinguistic skills. Consequently, the subjects' performance should give some kind of indication whether they had used linguistic and/or metalinguistic ready-made routines in their judgments.

It was assumed that if the subjects were slow, it might indicate a search in long-term memory (for relevant declarative knowledge). Certain types of subjects may be more inclined to use this strategy. It turned out that of the cognitive and affective variables included in the study only social and anxiety and, to a lesser degree, anxiety about errors appeared to be related to performance in RT judgment tasks. The subjects who felt more anxiety were either very accurate or very inaccurate. This is in line with what is known about the effect of anxiety: on the one hand, it may improve the performance of higher-ability learners; on the other hand, it may have a detrimental effect on poorer learners such as Group 3 in Experiment 3.

Another reason for longer RTs might also be L2 learners' lower levels of visual word recognition skills in L2. There appears to be some evidence for this in word recognition studies (eg. Segalowitz 1991). In addition, reading skills seem to play a great role in the development of metalinguistic skill (see eg. Bialystok and Ryan 1985a, 1985b). Since the sentences used in the experiments were very short, there is no reason to assume that higher-level L2 reading skills (such as inferencing) might have influenced the subjects' performance. It is more likely that if reading skills have any influence on test performance, it must be at the lower level of word recognition. Intelligence might also be another factor in RT judgments; there has been a number of studies conducted on the relationship of choicce reaction time to intelligence (see eg. Jensen and Reed 1990, Neubauer 1990).

The RT judgments may indeed turn out to be useful in revealing undelying differences among advanced language learners (cf. Segalowitz and Segalowitz 1993). When compared with the native-language RT judgments, the difference is revealing.

Another area in which RT judgments appear to be useful is in the study of language transfer. There emerged consistent and sometimes striking differences in RT judgments among learners with different L1s.

11.4 Limitations and suggestions for further research

The RT judgment experiments conducted here have, of course, the typical limitations of laboratory research. Linguists working within discourse analysis and pragmatics often point out, with some justification, that the findings obtained by such research may be difficult to generalize to the actual language use in real-life situations (cf. Hulstijn 1997).

However, if one accepts the necessary restrictions of research of this type, there are still some questions that must be raised. Grammaticality judgments are not without problems, as has been discussed above. It was assumed that fast reaction times would be an indication of automatic processing. This is an assumption frequently made in studies using reaction time measurement; however, this in itself may be problematic. Segalowitz and Segalowitz (1993), for example, discuss the difficulty in distinguishing between automatic and speeded controlled processing, assuming, of course, a fundamental qualitative difference between the two (see also Hulstijn 1997). Hulstijn (1997: 138) also points out the difficulty of determining whether speeded judgment tasks can be said to tap implicit and explicit knowledge: fast judgments may not necessarily indicate the use of implicit knowledge. The use of other types of tests and/or subjects simultaneously, however, makes it possible for reseachers to feel more confident of their findings. If the focus of research is in on the development of automaticity, or acquisition of skill, it appears important to conduct longitudinal studies that track the progress of learners at various stages of development (Hulstijn 1997).

These experimens had a more modest goal: the starting point was the empirically observed difference in performance between the subjects who were perceived as fluent skilled language users (ie., native speakers) and not so fluent skilled language users (language learners) on RT judgments of grammatical acceptability. An assumption was made that the difference could be explained by the difference in language proficiency. At the same time it was acknowledged that other factors, for example various cognitive and affective variables might also affect the performance in such test. No strong links were found. However, these experiments only used a handful of such factors; furthermore, the nature of the constructs themselves (eg. field-dependence/independence) may be debatable, not to mention their operationalizations. Therefore, it is difficult to draw any firm conclusions based on these findings. As to the relationship between RT judgments and L2 proficiency operationalized by a specific L2 proficiency test, the caveats used in language testing apply: What did the instrument measure? How reliable and valid was it? What kind of language knowledge was needed in order to successfully take the test?

The test itself focused on learners' receptive written skills. Further research is suggested in which, for example, L2 learners' reaction time latencies to deviant and non-deviant sentences is compared with their performance in a communicative task. This would also test the suggestions that RT judgments are indicative of communicative performance and that there might be a connection between the second language metalinguistic judgments and language production.

BIBLIOGRAPHY

Abraham, R. 1983. Relationship between use of the strategy of monitoring and cognitive style. *Studies in Second Language Acquisition*, 6, 17-32.

Abraham, R. 1985. Field independence-dependence and the teaching of grammar. *TESOL Quarterly*, 20, 689-702.

Adams, M. 1980. Failures to comprehend and levels of processing in reading. In R.J. Spiro, B.C. Bruce and W.F. Brewer (eds.), *Theoretical issues in reading comprehension*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 11-32.

Adjemian, C. 1976. On the nature of interlanguage systems. *Language Learning*, 26, 297-320.

Ahukanna, J.G.W., N.J. Lund and J.R. Gentile 1981. Inter- and intralingual interference effects when learning a third language. *Modern Languge Journal*, 65, 281-287.

Alanen, R. 1987. The influence of some features of English as the learner's L2 on his German: a study in cross-linguistic influence. An unpublished Pro Gradu Thesis. University of Jyväskylä.

Alanen, R. 1991. Cross-linguistic influence between native and non-native languages. An unpublished Licentiate Thesis. University of Jyväskylä.

Alanen, R. 1995. Input enhancement and rule presentation in second language acquisition. In R. Schmidt (ed.), *Attention and awareness in foreign language learning*. Second language teaching and curriculum center. University of Hawai'i at Manoa, 259-302.

Andersen, P.A. 1986. Consciousness, cognition, and communication, *The Western Journal of Speech Communication*, 50, 87-101.

Andersen, R.W. 1983. Transfer to somewhere. In S. Gass and L. Selinker (eds.), *Language transfer in language learning*. Rowley, Mass: Newbury House, 177-201.

Anderson, J.R. 1976. *Language, memory, and thought*. Hillsdale: New Jersey: Lawrence Erlbaum Associates.

Anderson, J.R.1980. *Cognitive psychology and implications*. San Francisco: Freeman.

Anderson, J.R. 1983. *The architecture of cognition*. Cambridge. Mass.:Harvard University Press.

d'Anglejean, A. and C. Renaud 1985. Learner characteristics and second language acquisition: a multivariate study of adult immigrants and some thoughts on methodology. *Language Learning*, 35, 1-19.

Arabski, J. 1968. A linguistic analysis of English composition errors made by Polish students. *Studia Anglica Posnaniensia*, 1, 71-89.

Ard, J. and S. Gass 1987. Lexical constraints on syntactic acquisition. *Studies in Second Language Acquisition*, 9, 233-251.

Ard, J. and T. Homburg 1983. Verification of languge transfer. In S. Gass and L. Selinker (eds.), *Language transfer in language learning*. Rowley, Mass: Newbury House, 157-176.

Arthur, B. 1980. Gauging the boundaries of second language competence: a study of learner judgments. *Language Learning*, 30, 177-194.

Ayoun, D. 1996. The subset principle in second language acquisition. *Applied Psycholinguistics*, 17, 185-213.

Bachman, L. 1990 *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Bachman, L. 1991 What does language testing have to offer? *TESOL Quarterly*, 25, 671-704.

Bachman, L. and A. Palmer 1984. Some comments on the terminology of language testing. In C. Rivera (ed.), *Communicative competence approaches to language proficiency: research and application*. Clevedon: Multilingual Matters, 34-43.

Bachman, L. and A. Palmer 1996. *Language testing in practice: designing and developing useful language tests*. Oxford: Oxford University Press.

Bates, E., B. MacWhinney, C. Caselli, A. Devescovi, F. Natale and V. Venza 1984. A cross-linguistic study of the development of sentence interpretation strategies. *Child Development*, 55, 341-354.

Bates, E. and B. MacWhinney 1979. A functionalist approach to the acquisition of grammar. In E. Ochs and B. Schieffelin (eds.), *Developmental pragmatics*. New York: Academic Press.

Bates, E. and B. MacWhinney 1982. Functionalist approaches to grammar. In E. Wanner and L. Gleitman (eds.), Language acquisition: the state of the art. New York: Cambridge University Press.

Bates, E. and B. MacWhinney 1989. Functionalism and the Competition Model, in B. MacWhinney and E. Bates (eds.), *The cross-linguistic study of sentence processing*. Cambridge: Cambridge University Press, 3-73

Bennink, C.D. 1982. Individual differences in cognitive style, working memory, and semantic integration. *Journal of Research in Personality*, 16, 267-280.

Berkoff, N.A. 1982. Error analysis revisited. In G. Nickel and D. Nehls (eds.), *Error analysis, contrastive linguistics and second language learning: papers from the 6th International Congress of Applied Linguistics, Lund 1981*, Special issue of IRAL. Heidelberg: Julius Groos Verlag, 5-17.

Bialystok, E. 1978. A theoretical model of second language learning. *Language Learning*, 28, 69-83.

Bialystok, E. 1979. Explicit and implicit judgments of L2 grammaticality. *Language learning*, 29, 81-103.

Bialystok, E. 1981. The role of linguistic knowledge in second language use. *Studies in Second Language Acquisition*, 4, 31-45.

Bialystok, E. 1982. On the relationship between knowing and using forms. *Applied Linguistics*, 3, 181-206.

Bialystok, E. 1990a. *Communication Strategies: a psychological analysis of second-language use*. Oxford: Blackwell.

Bialystok, E. 1990b. The dangers of dichotomy: a reply to Hulstijn. *Applied Linguistics*, 11, 46-51.

Bialystok, E. 1991. Achieving proficiency in a second language: a processing description. In R. Phillipson, E. Kellerman, L. Selinker, M . Sharwood Smith and M. Swain (eds.), *Foreign/Second language pedagogy research*. Clevedon: Multilingual Matters.

Bialystok, E. and E. Bouchard Ryan 1985a. A metacognitive framework for the development of first and second language skills. In D.L. Forrest-Presley, G.E. MacKinnon and T.G. Waller (eds.), *Metacognition, cognition, and human performance*. Orlando: Academic Press, 207-252.

Bialystok, E. and E. Bouchard Ryan 1985b. Toward a definition of metalinguistic skill. *Merrill-Palmer Quarterly*, 31, 229-251.

Bialystok, E. and M. Frohlich 1978. Variables of classroom achievement in second language learning. *Modern Language Journal*, 62, 327-336.

Bialystok, E. and M. Sharwood Smith 1985 Interlanguage is not a state of mind: an evaluation of the construct for second language acquisition. *Applied Linguistics*, 6, 101-117.

Birdsong, D. 1984. *Variables in variation*. Paper presented to the Foreign Language Education Center Student Association, University of Texas, Austin.

Birdsong, D. 1989a. *Metalinguistic performance and interlinguistic competence*. Berlin: Springer Verlag.

Birdsong, D. 1989b. Lux et verisimilitudo: judgment data in SLA theory. Paper presented at the Boston University Conference on Language Development (October). ERIC ED 312879.

Blackwell, A., Bates, E. and D. Fisher 1996. The time course of grammaticality judgment. *Language and Cognitive Processes*, 11, 337-406.

Blanco, M.J. and A.A. Alvarez 1994. Psychometric intelligence and visual focused attention: relationships in nonsearch tasks. *Intelligence*, 18, 77-106.

Bley-Vroman, R. 1989. The logical problem of second language learning. In S. Gass and J. Schachter (eds.), *Linguistic perspectives on second language acquisition*. Cambridge: Cambridge University Press, 41-68.

Bley-Vroman, R., Felix, S. and G. Ioup 1988. The accessibility of Universal Grammar in adult language learning. *Second Language Research*, 4, 1-32.

Bley-Vroman, R. and D. Masterson 1989. Reaction time as a supplement to grammaticality judgment in L2 studies. *University of Hawai'i Working Papers in ESL*, 8, 207-237.

Bourne, L.E., R.L. Dominowski, E.F. Loftus and A. F. Healy 1986. *Cognitive processes. Second edition*. Englewood Cliffs, NJ: Prentice-Hall.

Broadbent, D.E. 1981. Perceptual experiments and language theories. In H.C. Longuet-Higgins, J. Lyons and D.E. Broadbent (eds.) *The psychological mechanisms of language*, a joint symposium of the Royal Society and the British Academy, London: The Royal Society and the British Academy, 161-171.

Brown, G., K. Malmkjaer and J. Williams (eds.) 1996. *Performance and competence in second language acquisition*. Cambridge: Cambridge University Press.

Burgschmidt, E. and D. Götz 1974. *Kontrastive Linguistik Deutsch/ Englisch. Theorie und Anwendung*. München: Max Hueber.

Butterworth, B. 1980. Evidence from pauses in speech. In B. Butterworth (ed.), *Language production, Volume 1: Speech and talk*. London: Academic Press, 155-176.

Canale, M. and M. Swain 1980. Theoretical bases of communicative approaches to secodn language teaching and testing. *Applied Linguistics*, 1, 1-47.

Canale, M. 1983a. From communicative competence to communicative language pedagogy. In J.C. Richards and R.W. Schmidt (eds.), *Language and communication*. London: Longman, 2-27.

Canale, M. 1983b On some dimensions of language proficiency. In J.W. Oller (ed.), *Issues in language testing research*. Rowley, MA: Newbury House, 333-342.

Carlson, R.A., M. Sullivan and W. Schneider 1989. Practice and working memory effects in building procedural skill. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15, 517-526.

Carroll, J. 1981. Twenty-five years of research on foreign language aptitude. In K. Diller (ed.), *Individual differences and universals in language learning aptitude*. Rowley, MA: Newbury House.

Carter, E. 1988. The relationship of field-dependent/independent cognitive style to Spanish language achievement and proficiency: a preliminary report. *Modern Language Journal*, 72, 21-30.

Chapelle, C. and C. Roberts 1986. Ambiguity tolerance and field independence as predictors of proficiency in English as a second language. *Language Learning*, 36, 27-45.

Chapelle, C. and P. Green 1992. Field independence/dependence in second language acquisition research. *Language Learning* 42: 47-83.

Chaudron, C. 1983. Research on metalinguistic judgments: a review of theory, methods and results. *Language Learning*, 33, 343-377.

Chaudron, C. 1985. A method for examining the input/intake distinction. In S. Gass and C. Madden (eds.), *Input in second language acquisition*. Rowley, MA: Newbury House.

Chomsky, N. 1965. *Aspects of the theory of syntax*. Cambridge. Mass.: The MIT Press.

Chomsky, N. 1980. *Rules and representations*. Oxford: Blackwell.

Chomsky, N. 1986. *Knowledge of language: its nature, origin and use*. New York: Praeger.

Clahsen, H., J. Meisel and M. Pienemann 1983. *Deutsch als Zweitsprache: der Spracherwerb ausländischer Arbeiter*. Tübingen: Gunter Narr.

Clahsen, H. 1987. Connecting theories of language procesisng and (second) language acquisition. In C. Pfaff (ed.), *First and second language acquisition processes*. Cambridge: Newbury House.

Clark, E.V. 1978. Awareness of language: some evidence from what children say and do. In A.Sinclair, R.J.Jarvella and W.J.M. Levelt (eds.), *The child's conception of language*. Berlin: Springer-Verlag, 17-43.

Clark, H. and E. Clark 1977. *Psychology and language: an introduction to psycholinguistics*. New York: Harcourt Brace and Jovanovich.

Clément, R., R.C. Gardner and P.C. Smythe 1980. Social and individual factors in second language acquisition. *Canadian Journal of Behavioral Science*, 12, 293-302.

Cohen, J.D., K. Dunbar and J.L. McClelland 1990. On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychological Review*, 97, 332-361.

Cook, V. 1993. *Linguistics and second language acquisition*. London: MacMillan.

Cook, V. and M. Newson 1996. *Chomsky's Universal Grammar. An Introduction. Second edition.* Oxford: Blackwell.

Corder, S. Pit 1967. The significance of learners' errors. *International Review of Applied Linguistics*, 5, 161-169.

Coppieters, R. 1987. Competence differences between native and near-native speakers. *Language*, 63, 544-573.

Cowan, R. and Y. Hatasa 1994. Investigating the validity and reliability of native speaker and second-language learner judgments about sentences. In E. Tarone, S. Gass and A. Cohen (eds.), *Research methodology in second-language acquisition*. Hillsdale, NJ: Lawrence Erlbaum, 287-302

Coppieters, R. 1987. Competence differences between native and near-native speakers. *Language*, 63, 544-573.

Cruttenden, A. 1981. Item learning and system-learning. *Journal of Psycholinguistic Research*, 10, 79-88.

Cutler, A. and D. Norris 1979. Monitoring sentence comprehension: a study of the subjective lexicon. In W.E. Cooper and E.C.T. Walker (eds.), *Sentence processing:*

psycholinguistic studies presented to Merrill Garrett. Hillsdale, NJ: Lawrence Erlbaum.

Dagut, M. and B. Laufer 1982. How intralingual are 'Intralingual errors'?. In G. Nickel and D. Nehls (eds.), *Error analysis, contrastive linguistics and second language learning: papers from the 6th International Congress of Applied Linguistics, Lund 1981*, Special issue of IRAL, Heidelberg: Julius Groos Verlag, 19-42.

Danchev, A. 1982. Transfer and translation. *Finlance*, 2, 39-62.

Dechert, H.W., D. Möhle and M. Raupach (eds.) 1984. Second language productions. Tübingen: Gunter Narr.

Dechert, H.W. 1987. Understanding producing. In H. Dechert and M. Raupach (eds.), *Psycholinguistic models of production.* Norwood, NJ: Ablex, 229-234.

DeKeyser, R. 1994. Implicit and explicit learning of L2 grammar: a pilot study. *TESOL Quarterly*, 28, 188-194.

DeKeyser, R. 1996. Exploring automatization processes. *TESOL Quarterly*, 30, 349-357.

Dulany, D.E., R.A. Carlson and G.I. Dewey 1984. A case of syntactical learning and judgment: how conscious and how abstract? *Journal of Experimental Psychology: General*, 113, 541-555.

Dulay, H.C. and M. Burt 1974. Natural sequencies in acquisition. *Language Learning*, 24, 37-53.

Dulay, H., M. Burt and S. Krashen 1982. *Language two.* New York: Oxford University Press.

Eckman, F.R. 1977. Markedness and contrtastive analysis hypothesis. *Language Learning*, 27, 315-330.

Eckman, F.R. 1985. Some theoretical and pedagogical implications of the markedness differential hypothesis. *Studies in Second Language Acquisition*, 7, 289-307.

Eckman, F.R. 1994. The competence-performance issue in second-language acquisition theory: a debate. In E. Tarone, S. Gass and A. Cohen (eds.), *Research methodology in second-language acquisition.* Hillsdale, NJ: Lawrence Erlbaum, 3-15.

Ellis, N. 1994. Consciousness in second language learning: psychological perspectives pn the role of conscious processes in vocabulary acquisition. AILA Review,11, 37-56.

Ellis, N. and R. Schmidt 1997. Morphology and longer distance dependencies. Laboratory research illuminating the A in SLA. *Studies in Second Language Acquisition*, 19, 145-171.

Ellis, R. 1986. *Understanding second language acquisition.* Oxford: Oxford University Press.

Ellis, R. 1990. Individual learning styles in classroom second language development. In J. de Jong and D. Stevenson (eds.), *Individualizing the assessment of language liabilities.* Clevedon, Avon: Multilingual Matters.

Ellis, R. 1991. Grammaticality judgments and second language acquisition. *Studies in Second Language Acquisition*, 13, 161-186.

Ellis, R. 1994. *The study of second language acquisition.* Oxford: Oxford University Press.

Ely, C.M. 1986. An analysis of discomfort, risktaking, sociability, and motivation in the L2 classroom. *Language Learning*, 36, 1-25.

Enkvist, N. Success concepts. *Nordic research on text and discourse*. NORDTEXT Symposium (Espoo, Finland, May 10-13, 1990).

Ericsson, K. and H. Simon 1984. *Protocol analysis: verbal reports as data.* Cambridge, MA: MIT Press.

Eubank, L. (ed.) 1991. *Point Counterpoint. Universal Grammar in the second language.* Amsterdam: John Benjamins.

Faerch, C., K. Haastrup and R. Phillipson 1984. *Learner language and language learning.* Copenhagen: Gyldendals Sprogbibliotek.

Faerch, C. and G. Kasper 1983. *Strategies for interlanguage communication.* London: Longman.

Faerch, C. and G. Kasper 1984. Pragmatic knowledge: rules and procedures. *Applied Linguistics*, 5, 214-225.

Faerch, C. and G. Kasper 1986. The role of comprehension in second-language learning. *Applied Linguistics*, 7, 257-274.

Faerch, C. and G. Kasper 1987. Perspectives on language transfer. *Applied Linguistics*, 8, 111-136.

Favreau, M. and N. Segalowitz 1983. Automatic and controlled processes in the first and second language reading of fluent bilinguals. *Memory and Cognition*, 11, 565-574.

Flores d'Arcais, G. 1988. Language perception. In F.J. Newmeyer (ed.), *Linguistics: The Cambridge Survey. Volume III Language: Psychological and biological aspects*, Cambridge: Cambridge University Press, 97-123.

Fodor, J.A., T. Bever and M. Garrett 1974. *The psychology of language: an introduction to psycholinguistics and generative grammar.* New York: McGraw-Hill.

Forster, K. 1979. Levels of processing and the structure of the language processor. In W.E. Cooper and E.C.T. Walker (eds.), *Sentence processing: psycholinguistic studies presented to Merrill Garrett.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, 27-85.

Frazier, L. 1988. Grammar and language processing. In F.J. Newmeyer (ed.), *Linguistics: The Cambridge Survey II Linguistic theory: Extensions and implications*, Cambridge: Cambridge University Press, 15-34.

Forster, B. 1968. *The changing English language.* London: Macmillan.

Freedman, S. and K.I. Forster 1985. The psychological status of overgenerated sentences. *Cognition*, 19, 101-131.

Furth, H.G. and N.A. Milgram, Labeling and grouping effects in the recall of pictures by children. *Child Development*, 44, 511-518.

Galambos, S.J. and S. Goldin-Meadow 1986. A study of the development of metalinguistic awareness in bilingual and monolingual children. Unpublished MS: Department of Psychology, Yale University, New Haven.

Galambos, S.J. and K. Hakuta 1988. Subject-specific and task-specific characteristics of metalinguistic awareness in bilingual children. *Applied Psycholinguistics*, 9, 141-162.

Gardner, R.C. 1985. *Social psychology and second language learning: the role of attitudes and motivation.* London: Edward Arnold.

Gass, S. 1983a. Language transfer and universal grammatical relations. In S. Gass and L. Selinker (eds.) 1983, *Language transfer in language learning.* Rowley, Mass: Newbury House, 69-82.

Gass, S. 1983b. The development of L2 intuitions. *TESOL Quarterly*, 17, 273-291.

Gass, S. 1988. Integrating research areas: a framework for second language studies. *Applied Linguistics*, 9, 198-217.

Gass, S. 1994. The reliability of second-language grammaticality judgments. In E. Tarone, S. Gass and A. Cohen (eds.), Research methodology in second-language acquisition. Hillsdale, NJ: Lawrence Erlbaum, 303-322.

Gass, S. and L. Selinker 1983. Introduction. In S. Gass and L. Selinker (eds.), *Language transfer in language learning*. Rowley, Mass: Newbury House, 1-18.

Genesee, F. 1988. Neuropsychology and second language acquisition, in L.M. Beebe (ed.), *Issues in second language acquisition*. New York: Newbury House, 81-112.

Genesee, F. and E. Hamayan 1980. Individual differences in second language learning. *Applied Psycholinguistics*, 1, 95-110.

Grauberg, N. 1971. An error analysis in German of first-year university students, in G.E. Perren and J.J.M. Trimm (eds.), *Applications of Linguistics*. Cambridge: Cambridge University Press.

Greenbaum, S. 1973. Informant elicitation of data on syntactic variation. *Lingua*, 31, 201-212.

Gregg, K. 1984. Krashen's Monitor and Occam's Razor. *Applied Linguistics*, 5, 79-100.

Gregg, K. 1990. The variable competence model of second language acquisition, and why it isn't. *Applied Linguistics*, 11, 365-383.

Griffiths, R. and R. Sheen 1992. Disembedded figures in the landscape: a reappraisal of L2 research on field dependence/independence. *Applied Linguistics*, 13, 133-148.

Hakulinen, L. 1979. *Suomen kielen rakenne ja kehitys*. Neljäs, korjattu ja lisätty painos. Helsinki: Otava.

Hakulinen, A. and F. Karlsson 1979. *Nykysuomen lauseoppia*. Jyväskylä: Gummerus, Suomalaisen kirjallisuuden Seura.

Halliday, M. 1978. *Language as a social semiotic*. London: Edward Arnold.

Hansen, J. and C. Stansfield 1981. The relationship of field dependent-independent cognitive styles to foreign language achievement. *Language Learning*, 31, 349-367.

Hansen, L. 1984. Field dependence-independence and language testing: evidence from six Pacific Island culutures. *TESOL Quarterly*, 18, 311-324.

Harrington, M. 1987. Processing transfer: language-specific strategies as a source of interlanguage variation. *Applied Psycholinguistics*, 8, 351-378.

Hatch E. and H. Farhady 1982. *Research design and statistics for applied linguistics*. Rowley, Mass.: Newbury House.

Havia, V. 1982. *Language processing by Finnish and native speakers of English with special reference to the use of grammatical knowledge*. Un unpublished Pro Gradu Thesis. University of Jyväskylä.

Hedgcock, J. 1992. L2 metalingual judgments: why do learners respond differently to grammatical vs. ungrammatical strings? Paper presented at the Annual Meeting of teh American Association for Applied Linguistics, Seattle, WA (March).

Hedgcock, J. 1993. Well-formed vs. ill-formed strings in L2 metalingual tasks: specifying features of grammaticality judgments. *Second Language Research*, 9, 1-21.

Hedgcock, J. and N. Lefkowitz 1993. Tuning in on the "Prime Time": channel effects and priming in L2 grammaticality judgments. *Foreign Language Annals*, 26, 297-307.

Helbig, G. and J. Buscha. 1984. *Deutsche Grammatik: ein Handbuch für den Ausländerunterricht*. Leipzig: VEB Verlag Enzyklopädie Leipzig.

Hellinger, M. 1977. *Kontrastive Grammatik Deutsch/Englisch*. Anglistische Arbeitshefte 14. Tübingen: Max Niemayer.

Herranen, T. 1977. *Errors made by Finnsih university students in the use of the English article system*. Un unpublished Pro Gradu Thesis. University of Jyväskylä.

Hirakawa, M. 1989. The governing category parameter in second language acquisition. *MITA Working Papers in Psycholinguistics*, 2, 1-15.

Hopkins, E.A. 1982. Contrastive analysis, interlanguage, and the learner. In W.F.W. Lohnes and E.A. Hopkins (eds.), *The contrastive grammar of English and German*. Ann Arbor, Michigan: Karoma.

Horwitz, E.K., M.B. Horwitz and J. Cope 1986. Foreign language classroom anxiety. *Modern Language Journal*, 70, 125-132.

Hulstijn, J.H. 1982. *Monitor use by adult second language learners.* Unpublished dissertation. Krips Repro Meppel. Amsterdam.

Hulstijn, J. 1989. Implicit and incidental second language learning: experiments in the processing of natural and partly aritificial input. In H. Dechert and M. Raupach (eds.), *Interlingual processes*. Tübingen: Gunter Narr,

Hulstijn, J. 1990. A comparison between the information-processing and the analysis/control approaches to language learning. *Applied Linguistics*, 11, 30-45.

Hulstijn, J. 1997. Second language acquisition research in the laboratory. Possibilities and limitations. *Studies in Second Language Acquisition*, 19, 131-143.

Hulstijn, J. and W. Hulstijn 1984. Grammatical errors as a function of processing constraints and explicit knowledge. *Language Learning*, 34, 23-43.

Hyltenstam, K. 1983. Data Types and Second Language Variability. In H. Ringbom (ed.), *Psycholinguistics and Foreign Language Learning. Papers from a conference (Stockholm, Sweden and Abo, Finland, October 25-26, 1982)*.

Inaba, M. 1993. Examining the including and excluding roles of positive evidence: a study of a case where L2/L1 grammar intersects. *MITA Working Papers in Psycholinguistics*, 3, 33-54.

James, C. 1980. *Contrastive analysis.* London: Longman.

Jamieson, J. 1992. The cognitive styles of reflection/impulsivity and field independence/dependence and ESL success. *Modern Language Journal*, 76, 491-501.

Jensen, A.R. 1985 Methodological and statistical techniques for the chronometric study of mental abilities. In C.R. Reynolds & V. Willson (eds.), *Methodological and statistical advances in the study of individual differences*. New York: Plenum Press.

Jensen, A.R. and T.E. Reed 1990. Simple reaction time as a suppressor variable in the chronometric study of intelligence. *Intelligence*, 14, 375-388.

Johnson, K. 1996. *Language teaching and skill learning*. Oxford: Blackwell.

Johnson, J.S. and E. Newport 1989. Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21, 60-99.

Jordens, P. 1977. Rules, grammatical intuitions ans strategies in foreign language learning. *Interlanguage Studies Bulletin*, 2.2, 5-76.

Juffs, A. and M. Harrington, 1995. Parsing effects in second language sentence processing: subject and object asymmetries in "Wh"-extraction. *Studies in Second Language Acquisition*, 17, 483-516.

Just, M.A. and P.A. Carpenter 1992. A capacity theory of comprehension: individual differences in working memory. *Psychological Review*, 99, 122-149.

Juhasz, J. 1970. Kontrastive und Interferenzaspekte der Sprache. Interferenzlinguistik. In H. Althaus, H. Henne and H. Wiegand (eds.), *Lexikon der germanistischen Linguistik*. Tübingen: Max Niemeyer.

Jung, U.O.H. 1981. Englisch als Fremdsprache für Kinder mit Deutsch als Zweitsprache. *Grazer Linguistische Studien*, 14, 83-97.

Kahneman, D. 1973. *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.

Karlsson, F. 1979. Automatic morphological segmentation of Finnish word forms. In K. Häkkinen and F. Karlsson (eds.), *Papers from the conference on general linguistics, Seili 29.-30.8.1979.* Publications of the Linguistic Association of Finland, 3, Turku: The Linguistic Association of Finland, 51-64.

Karlsson, F. 1983. *Finnish grammar.* Porvoo: WSOY.

Karmiloff-Smith, A. 1986a. Stage/structure versus phase/process in modelling linguistic and cognitive development. In I. Levin (ed.), Stage and structure: reopening the debate. Norwood, NJ: Ablex, 164-191.

Karmiloff-Smith, A. 1986b. From meta-processes to conscious access: evidence from children's metalinguistic and repair data. *Cognition,* 23, 95-147.

Kellerman, E. 1977. Towards a characterisation of the strategy of transfer in second language learning. *Interlanguage Studies Bulletin,* 2.1, 58-145.

Kellerman, E. 1978. Giving learners a break: native language intuitions as a source of predictions about transferability. *Working Papers on Bilingualism,* 15, 59-92.

Kellerman, E. 1983. Now you see it, now you don't, in S. Gass and L. Selinker (eds.) *Language transfer in language learning.* Rowley, Mass: Newbury House Publishers, 112-134.

Kilborn, K. and A. Cooreman 1987. Sentence interpretation strategies in adult Dutch-English bilinguals. *Applied Psycholinguistics,* 8, 415-431.

Kilborn, K. 1989 Sentence processing in a second language: the timing of transfer. *Language and Speech,* 32, 1-23.

Kim, R. 1993. A sensitive period for second language acquisition: a reaction-time grammaticality judgment task with Korean-English bilinguals. IDEAL, 6, 15-27.

Klein, W. 1986. *Second language acquisition.* Cambridge: Cambridge University Press.

Knapp-Potthoff, A. and K. Knapp 1982. *Fremdsprachenlernen und -lehren. Eine Einführung in die Didaktik der Fremdsprachen vom Standpunkt der Zweitsprachenerwerbsforschung.* Stuttgart: W. Kohlhammer.

Koponen, M. 1992. Glottal boundary markers and aspects of disfluency in Finnish English and epilogue (from 1990- 1992). An unpublished Licentiate Thesis. University of Jyväskylä.

Korhonen, O. 1983. *Erroneous use of English verb forms by Finnish schoolchildren.* Un unpublished Pro Gradu Thesis in English Philology. University of Jyväskylä.

Krashen, S. 1978. The Monitor Model for second-language acquisition. In R.C. Gingras (ed.), Second language acquisition and foreign language teaching. Arlington, VA: Center for Applied Linguistics.

Krashen, S. 1981. *Second language acquisition and second language learning.* Oxford: Pergamon Press.

Krashen, S. 1982. *Principles and practice in second language acquisition.* Oxford: Pergamon Press.

Krashen, S. 1983. Newmark's 'Ignorance Hypothesis' and current second language acquisition theory. In S. Gass and L. Selinker (eds.), *Language transfer in language learning,* Rowley, Mass: Newbury House Publishers, 135-153.

Lachman, R., J.L. Lachman and E.C. Butterfield 1979. *Cognitive psychology and information processing: an introduction.* Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Lado, R. 1957. *Linguistic across Cultures.* Ann Arbor, Michigan: University of Michigan Press.

Lamminpää, S. 1980. *Production strategies in the translations from Finnish into English by Finnish students of English. A study of lexical errors.* An unpublished Pro Gradu Thesis. University of Jyväskylä.

Lehtonen, J. 1982. The theory and methodology of speech science and contrastive analysis. *Papers and Studies in Contrastive Linguistics*, 14. Poznan: Adam Mickiewicz University, 45-58.

Lehtonen, J. 1983. Orthography as a source of the Finns' difficulties in foreign-language speech processing, in C-C. Elert and A. Seppänen (eds.), *Finnish-English language contact: papers from a workshop, May 4, 1981. Umeå Papers in English, 4.* Umeå: University of Umeå, 5-22.

Lehtonen, J. 1984. Reaktioaikamittaukset psykolingvistiikan tutkimusmenetelmänä. In K. Sajavaara, J. Tommola and M. Leiwo (eds.), *Psycholinguistic Papers IV.* Publications of the Finnish Association of Applied Linguistics 37, Jyväskylä: AFinLA, 75-85.

Lehtonen, J. and K. Sajavaara 1983. Acceptability and ambiguity in native and second language message processing. In H. Ringbom (ed.), *Psycholinguistics and foreign language learning.* Publications of the Research Institute of the Åbo Akademi Foundation 86, Turku: Åbo Akademi, 101-125.

Lehtonen, J. and K. Sajavaara 1984. Phonology and speech processing in cross-language communication. In S. Eliasson (ed.). *Theoretical issues in contrastive phonology.* Heidelberg: Julius Groos, 85-99.

Lehtonen, J. and K. Sajavaara 1985. Psycholinguistic testing of transfer in foreign-language speech processing. *LAUT Series B* 117. Trier: University of Trier.

Lehtonen, J., K. Sajavaara and S. Manninen 1985. Communication apprehension and attitudes towards a foreign language. *Scandinavian Working Papers on Bilingualism,* 5, 53-62.

Lehtonen, J., K. Sajavaara and A. May 1976. Spoken English.

Leontev, A.A. 1973. Some problems in learning Russian as a foreign language. *Soviet Psychology, 11.*

Levelt, W.J.M., J.A.W.M. van Gent, A.F.J. Haans and A.J.A. Meijers 1977 Grammaticality, paraphrase and imagery. In S. Greenbaum (ed.), *Acceptability in language.* The Hague: Mouton, 87-101.

Levelt, W.J.M. 1978. A survey of studies in sentence perception: 1970-1976. In W.J.M. Levelt and G. Flores d'Arcais (eds.), *Studies in the perception of language.* Chichester: John Wiley & Sons, 1-74.

Levelt, W.J.M. 1989. *Speaking. From intention to articulation.* Cambridge, Mass.: The MIT Press.

LoCoco, V. 1976. A cross-sectional study on L3 acquisition. *Working Papers on Bilingualism,* 9, 44-75.

Logan, G. 1985. Skill and automaticity: relations, implications, and future directions, *Canadian Journal of Psychology,* 39, 367-386.

Logan, G.D. 1988. Toward an instance theory of automatization. *Psychological review,* 95, 492-527.

Logan, G.D. and S.T. Klapp 1991. Automatizing alphabet arithmetic problems 1: Is extended practice necessary to produce automaticity? *Journal of Experimental Psychology: Learning, Memory and Cognition,* 17, 179-195.

Logan, G.D. and J.L. Etherton 1994. What is learned during automatization? The role of attention in constructing an instance. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 20, 1022-1050.

Loftus, G.R., D. J. Miller and H.J. Burns 1978. Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory*, 4, 19-31.

Lukatela, K., S. Crain and D. Shankweiler 1988. Sensitivity to inflectional morphology in agrammatism: investigation of a highly inflected language. Brain and Language, 33, 1-15.

Lyons, J. 1969. *Introduction to theoretical linguistics*. Cambridge: Cambridge University Press.

MacDonald, M.C., N.J. Pearlmutter and M.S. Seidenberg 1994. Syntactic ambiguity resolution as lexical ambiguity resolution. In C. Clifton, L. Frazier and K. Rayner (eds.), *Perspectives on sentence processing*. Hillsdale, NJ: Lawrence Erlbaum, 123-153.

MacIntyre, P. and R. Gardner 1989. Anxiety and second language learning: toward a theoretical clarification. Language Learning, 39, 251-281.

MacIntyre, P. and R. Gardner 1991. Methods and results in the study of anxiety and language learning: a review of the literature. Language Learning, 41, 85-117.

MacIntyre, P. and R. Gardner 1994. The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning*, 44, 283-305.

MacWhinney, B., E. Bates and R. Kliegl 1984. Cue validity and sentence interpretation in English, German, and Italian. *Journal of verbal learning and verbal behavior*, 23, 127-150.

MacWhinney, B. 1987. Applying the Competition Model to bilingualism. *Applied Psycholinguistics*, 8, 315-327.

Marslen-Wilson, W.D. and A. Welsh 1978. Processing interaction and lexical access during word recognition in continuous speech. *Cognitive Psychology*, 10, 29-63.

Marslen-Wilson, W.D. and L.K. Tyler 1980. The temporal structure of spoken language understanding. *Cognition*, 8, 1-71.

Marslen-Wilson, W.D. 1987. Functional parallellism in spoken word-recognition. *Cognition*, 25, 71-102.

Masny, D. and A. d'Anglejean 1985. Second language grammaticality judgments. *Journal of Psycholinguistic Research, 14*, 175-197.

Mazurkewich, I. and L. White 1984. The acquisition of the dative alternation: unlearning overgeneralization. *Cognition*, 16, 261-283.

McClelland, J. and D.E. Rumelhart 1985. Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General*, 114, 159-188.

McCroskey, J.C. 1984. The communication apprehension perspective. In J.A. Daly and J.M. McCroskey (eds.), *Avoiding communication: shyness, reticence and communication apprehension*. Beverly Hills: Sage, 13-38.

McLaughlin, B. 1978. *Second-language acquisition in childhood*. Hillsdale, NJ: Lawrence Erlbaum.

McLaughlin, B. 1980. Theory and research in second-language learning: an emerging paradigm. *Language Learning*, 30, 331-350.

McDonald, J. 1987 Sentence interpretation in bilingual speakers of English and Dutch. *Applied Psycholinguistics*, 8, 379-413.

McLaughlin, B. 1987 *Theories of second-language learning*. London: Edward Arnold.

McLaughlin, B, 1990 "Conscious" versus "unconscious" learning. *TESOL Quarterly*, 24, 617-634.

McLaughlin, B. and N. Nayak 1989. Processing a new language: does knowing other languages make a difference? In H.W. Dechert and M. Raupach (eds.), *Interlingual processes*. Tübingen: Gunter Narr, 5-16.

McLaughlin, B. , T. Rossman and B. McLeod 1983. Second language learning: an information-processing perspective. *Language Learning*, 33, 135-158.

McLeod, B. and B. McLaughlin 1986. Restructuring or automaticity. *Language Learning*, 36, 107-156.

McNamara, T. 1996. *Measuring second language performance*. London: Longman.

Messick, S. 1976 Personality consistencies in cognition and creativity. In S. Messick (ed.), *Individuality in learning: cognitive styles and creativity for human development*. San Francisco: Jossey-Bass, 4-22.

Morrissey, M.D. 1983. Towards a grammar of learners' errors. *International Review of Applied Linguistics*, 21, 193-207.

Möhle, D. and M. Raupach 1989. Language transfer of procedural knowledge. In H.W. Dechert and M. Raupach (eds.), *Transfer in language production*. Norwood, NJ: Ablex, 195-216.

Mägiste, E. 1984. Learning a third language. *Journal of Multilingual and Multicultural Development*, 5, 415-421.

Nagata, H. 1987. Long-term effect of repetition on judgments of grammaticality.*Perceptual and Motor Skills*, 65, 295-299.

Nagata, H. 1988. The relativity of linguistic intuitions: the effect of repetition on grammaticality judgments. *Journal of Psycholinguistic Research*, 17, 1-17.

Nagata, H. 1989a. Judgments of sentence grammaticality and field-dependence of subjects. *Perceptual and Motor Skills*, 69, 739-747.

Nagata, H. 1989b. Effect of repetition on grammaticality judgments under objective and subjective self-awareness conditions. Journal of Psycholinguistic Research, 17, 1-17.

Nagata, H. 1989c. Repetition effect in judgments of grammaticality of sentences: examination with ungrammatical sentences. *Perceptual and Motor Skills*, 68, 275-282.

Naiman, N., M. Frohlich, H.H. Stern and A. Todesco 1978. *The good language learner*, Research in Education Series, 7. Ontario Institute for Studies in Education.

Nemser, W. 1971. Approximative systems of foreign language learners. *International Review of Applied Linguistics*, 9, 115-123.

Neubauer, A.C. 1990. Selective reaction times and intelligence. *Intelligence*, 14, 79-96.

Neubauer, A.C. and H.H. Freudenthaler 1994. Reaction times in a sentence-picture verification test and intelligence: individual strategies and effects of extended practice. *Intelligence*, 19, 193-218.

Odlin, T. 1989. *Language transfer*. Cambridge: Cambridge University Press.

Osterhout, L. 1994. Event-related brain potentials as tools for comprending language comprehension. In C. Clifton, L. Frazier and K. Rayner (eds.), *Perspectives on sentence processing*. Hillsdale, NJ: Lawrence Erlbaum, 15-44.

Paivio, A. 1971. *Imagery and verbal processes*. New York: Holt, Reinhart and Winston.

Paivio, A. 1978. The relationship between verbal and perceptual codes. In E.C. Carterette and M. Friedman (eds.), *Handbook of perception, Volume VIII. Perceptual coding.* New York: Academic Press, 375-397.

Paivio, A. and R. Harshman 1983. Factor analysis of a questionnaire on imagery and verbal habits and skills. *Canadian Journal of Psychology*, 37, 461-483.

Pawley, A. and F.H. Syder 1983. Two puzzles for linguistic theory: nativelike selection and nativelike fluency. In J.C. Richards and R.W. Schmidt (eds.), *Language and communication.* London: Longman, 191-225.

Perruchet, P. and C. Pacteau 1990. Synthetic grammar learning: implicit rule abstraction or explicit fragmentary knowledge? *Journal of Experimental Psychology: General*, 119, 264-275.

Pienemann, M. 1989. Is language teachable? *Applied Linguistics*, 10, 52-79.

Pinker, S. and A. Prince 1988. On language and connectionism: analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28, 73-193.

Posner, M. and C. Snyder 1975. Attention and cognitive control. In R. Solso (ed.), *Information processing and cognition: The Loyola symposium.* Hillsdale, NJ: Lawrence Erlbaum, 55-85.

Prideaux, G. 1984. *Psycholinguistics: the experimental study of language*. London: Croom Helm.

Pylyshyn, Z.W. 1973. What the mind's eye tells the mind's brain: a critique of mental imagery. *Psychological Bulletin*, 80, 1-24.

Quirk, R., S. Greenbaum, G. Leech and J. Svartvik 1973. *A grammar of contemporary English.* London: Longman.

Quirk, R. and J. Svartvik 1966. *Investigating linguistic acceptability*. The Hague: Mouton.

Reber, A. 1976. Implicit learning of synthetic languages: the role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory*, 2, 88-94.

Reber, A., R. Allen and S. Regan 1985. Sybtactic learning and judgments: still unconscious and still abstract. *Journal of Experimental Psychology: General*, 114, 17-24.

Ringbom, H. 1982. The influence of other languages on the vocabulary of foreign language learners. In G. Nickel and D. Nehls (eds.), *Error analysis, contrastive linguistics and second language learning: papers from the 6th International Congress of Applied Linguistics, Lund 1981*, Special issue of International Review of Applied Linguistics. Heidelberg: Julius Groos Verlag, 85-96.

Ringbom, H. 1983. Borrowing and lexical transfer. *Applied Linguistics*, 4, 207-212.

Ringbom, H. 1985a. Transfer in relation to some other variables in L2-learning. In H. Ringbom (ed.), *Foreign language learning and bilingualism.* Publications of the Research Institute of the Åbo Akademi Foundation, Turku: Åbo Akademi. 9-21.

Ringbom, H. 1985b. The influence of Swedish on the English of the Finnish learners, in H. Ringbom (ed.), *Foreign language learning and bilingualism*, Publications of the Research Institute of the Åbo Akademi Foundation, Turku: Åbo Akademi, 9-21.

Ringbom, H. 1986. Cross-linguistic influence and the foreign language learning process. In E. Kellerman and M. Sharwood Smith (eds.), *Crosslinguistic influence in second language acquisition.* New York: Pergamon Press, 150-162.

Ringbom, H. 1987. *The role of the first language in foreign language learning*. Clevedon: Multilingual Matters.

Robinson, P. 1994. Universals of word formation processes: noun incorporation in the acquisition of Samoan as a second language. *Language Learning*, 44, 569-601.

Robinson, P. and M.A. Ha 1993 Instance theory and second language rule learning under explicit conditions. *Studies in Second Language Acquisition*, 15, 413-438.

Rogers, M. 1987. Learners' difficulties with grammatical gender in German as a foreign language. *Applied Linguistics*, 8, 48-73.

Rutherford, W. 1988. *Second language grammar: learning and teaching*. London: Longman.

Ryan, E.B. and G.W. Ledger 1979. Grammaticality judgments, sentence repetitions, and sentence corrections of children learning to read. International Journal of Psycholinguistics, 6, 23-40.

Ryle, G. 1949. *The concept of mind*. London: Hutchinson.

Räsänen, A. 1976. *Variations in views of acceptability held by native speakers of English resident in Finland*. An unpublished Pro Gradu Thesis. University of Jyväskylä.

Sajavaara, K. 1981a. Psycholinguistic models, second language acquisition, and contrastive analysis. In J. Fisiak (ed.), *Contrastive linguistics and the language teacher*. Oxford: Pergamon Institute of English, 87-120.

Sajavaara, K. 1981b. Contrastive linguistics past and present and a communicative approach. In J. Fisiak (ed.), *Contrastive linguistics and the language teacher*. Oxford: Pergamon Institute of English, 33-56.

Sajavaara, K. 1983. The article errors of Finnish learners of English, In C-C. Elert and A. Seppänen (eds.), *Finnish-English language contact: papers from a workshop, May 4, 1981*. Umeå Papers in English 4. Umeå: University of Umeå, 72-87.

Sajavaara, K. 1987a. Intake in second language acquisition. In W. Lörscher and R. Schulze (eds.), *Perspectives on language in performance: studies in linguistics, literary criticism, and foreign language methodology: to honour Werner Hüllen on the occasion of his sixtieh birthday, 17 October 1987*. Tübingen: Gunter Narr.

Sajavaara, K. 1987b. Second language speech production: factors affecting fluency: In H. Dechert and M. Raupach (eds.), *Psycholinguistic models of production*. Norwood, NJ: Ablex, 45-75.

Sajavaara, K. and J. Lehtonen 1980. Language teaching and acquisition of communication. In K. Sajavaara, A. Räsänen and T. Hirvonen (eds.), *AFinLA yearbook*. Publications of the Finnish Association of Applied Linguistics 28, Jyväskylä: AFinLA, 25-35.

Sajavaara, K. and J. Lehtonen 1986. The mother tongue and the foreign language interaction. In D. Kastovsky and A. Szwedek (eds.) *Linguistics across historical and geographical boundaries: in honour of Jacek Fisiak on the occasion of his fiftieth birthday*. Trends in linguistics. Studies and Monographs 32, New York: Mouton de Gruyter.

Sajavaara, K. and J. Lehtonen 1989. Aspects of transfer in foreign language speakers' reactions to acceptability. In H.W. Dechert and M. Raupach (eds.), *Transfer in language production*. Norwood, NJ: Ablex, 35-52.

Schachter, J., A. Tyson, and F. Diffley 1976. Learner intuitions on grammaticality. *Language Learning*, 26, 67-76.

Schachter, J. 1974. An error in error analysis. *Language Learning* 24, 205-214.

Schachter, J. 1983. A new account of language transfer. In S. Gass and L. Selinker (eds.), *Language transfer in language learning*. Rowley, Mass.: Newbury House, 98-111.

Schachter, J. 1988. Second language acquisition and its relationship to Universal Grammar. *Applied Linguistics*, 9, 219-235.

Schachter, J. and V. Yip 1990. Grammaticality judgments: why does anyone object to subject extraction? *Studies in Second Language Acquisition*, 12, 379-392.

Schmidt, R. 1990 The role of consciousness in second language learning. *Applied Linguistics*, 11, 129-158.

Schmidt, R. 1992 Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition*, 14, 357-385.

Schmidt, R. 1994 Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review*, 11, 11-26.

Schmidt, R. and S. Frota 1986. Developing basic conversational ability in a second language: a case-study of an adult learner. In R. Day (ed.), *Talking to learn: conversation in second language acquisition.* Rowley, Mass: Newbury House.

Schneider, W. 1985. Toward a model of attention and the development of automatic processing. In M. Posner and O. Maris (eds.), *Attention and performance XI.* Hillsdale, NJ: Lawrence Erlbaum, 475-492.

Schneider, W. and M. Detweiler 1988. The role of practice in dual-task performance: Toward workload modeling in a connectionist/control architecture. *Human Factors*, 30, 539-566.

Schneider, W., S. Dumais and R. Shiffrin 1984. Automatic and control processing and attention. In R. Parasuraman and D. Davies (eds.), *Varieties of attention.* Orlando, FL: Academic Press, 1-27.

Segalowitz, N. S. 1986. Skilled reading in the second language. In J. Vaid (ed.), *Language processing in bilinguals: Psycholinguistic and neuropsychological perspectives.* Hilssdale, NJ: Lawrence Erlbaum, 3-19.

Segalowitz, N. S. 1991. Does advanced skill in a second language reduce automaticity in the first language? *Language Learning*, 41, 59-83.

Segalowitz, N.S. and S.J. Segalowitz 1993. Skilled performance, practice, and the differentiation of speed-up from automatization effects: evidence from second language word recognition. *Applied Psycholinguistics*, 14, 369-385.

Selinker, L. 1972. Interlanguage. *International Review of Applied Linguistics*, 10, 209-231.

Selinker, L. 1992. *Rediscovering interlanguage*. London: Longman.

Selinker, L., M. Swain and G. Dumas 1975. The interlanguage hypothesis extended to children. *Language Learning* 25, 139-191.

Sharwood Smith, M. 1981. Consciousness-raising and the second language learner. *Applied Linguistics*, 2, 159-169.

Sharwood Smith, M. 1982. Language transfer: the state of the art. *Finlance*, 2, 27-38.

Sharwood Smith, M. 1986. Comprehension versus acquisition: two ways of processing input. *Applied Linguistics* 3, 239-256.

Sharwood Smith, M. 1993. Input enhancement in instructed second language acquisition: theoretical bases. *Studies in Second Language Acquisition*, 15, 165-180.

Sharwood Smith, M. 1994 *Second language learning: theoretical foundations*. London: Longman.

Shiffrin, R.M. and W. Schneider 1977. Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, 84, 127-190.

Shim, R.K. 1993. Sensitive periods for second language acquisition: a reaction-time study of Korean-English bilinguals. *Issues and Developments in English and Applied Linguistics (IDEAL)*. Illinois University, Urbana. Division of English as an International Language.

Singh, R. and S. Carroll 1979. L1, L2 and L3. *Indian Journal of Linguistics* 5, 51-63.

Sjöholm, K. 1976. A comparison of the test results in grammar and vocabulary between Finnish- and Swedish-speaking applicants for English, 1974. In H. Ringbom and R. Palmberg (eds.), *Errors made by Finns and Swedish-speaking Finns in the learning of English*. AFTIL 5, Turku: Department of English, 54-137.

Skehan, P. 1986 Cluster analysis and the identification of learner types. In V. Cook (ed.), *Experimental approaches to second language acquisition*. Oxford: Pergamon, 81-94.

Skehan, P. 1989. *Individual differences in second-language learning*. London: Edward Arnold.

Slobin, D. 1973. Cognitive Prerequisites for the development of grammar. In C. Ferguson and D. Slobin (eds.), *Studies in child language development*. New York: Holt, Rinehart and Winston, 175-208.

Slobin, D. 1979. *Psycholinguistics*. Second edition. Glenview, Illinois: Scott, Freeman and Company.

Slobin, D. 1985. Cross-linguistic evidence for the language-making capacity. In D. Slobin (ed.), *The cross-linguistic study of language acquisition. Vol. 2: Theoretical issues*. Hillsdale, NJ: Lawrence Erlbaum.

Spitzer, S. E. 1986. *The measurement and implications of visual and verbal cognitive styles*. Ann Arbor, Michigan: UMI.

Stansfield, C. and J. Hansen 1983. Field-dependence-independence as a variable in second language cloze test performance. *TESOL Quarterly*, 17, 29-38.

Stedje, A. Tredjespråksinterferens i fritt tal - en jämförande studie. In R. Palmberg and H. Ringbom (eds.), *Föredrag vid konferensen om kontrastiv lingvistik och felanalys, Stockholm and Åbo, 7-8 februari, 1977*. Turku: Åbo Akademi, 141-158.

Sutter, J.C. and C.J. Johnson 1990. School-age children's metalinguistic awareness of grammaticality in verb form. *Journal of Speech and Hearing Research*, 33, 84-95.

Steinberg, D. 1982. *Psycholinguistics: language, mind, and world*. London: Longman.

Taft, M. 1986. Lexical access codes in visual and auditory word recognition. *Language and Cognitive Processes*, 1, 297-308.

Tarvainen,K. 1985. *Kontrastive Syntax Deutsch-Finnisch*. Deutsch im Kontrast, Band 4. Heidelberg: Julius Groos Verlag.

Taskinen, A. 1981. *Errors by Finnish learners of English and universal errors*. An unpublished Pro Gradu Thesis. University of Jyväskylä.

Taylor, R. 1975. The use of overgeneralization and transfer learning strategies by elementary and intermediate students of ESL. *Language Learning*, 25, 73-107.

Tobias, S. 1986. Anxiety and cognitive processing of instruction. In R. Schwartzer (ed.), *Self related cognition in anxiety and motivation*. Hillsdale, NJ: Lawrence Erlbaum, 33-54.

Tomasello, M. and C. Herron 1989. Feedback for language transfer errors: the Garden Path Technique. *Studies in Second Language Acquisition*, 11, 385-395.

Toury, G. 1986. Monitoring discourse transfer: a test-case for a developmental model of translation. In J. House and S. Blum-Kulka (eds.), *Interlingual and intercultural communication: discourse and cognition in translation and second language acquisition studies*. Tübingen: Gunter Narr, 79-94.

Towell, R., R. Hawkins and N. Bazergui 1993. Systemic and nonsystemic variability in advanced language learning. *Studies in Second Language Acquisition*, 15, 439-460.

Towell, R., R. Hawkins and N. Bazergui 1996. The development of fluency in advanced learners of French. *Applied Linguistics*, 17, 84-115.

Tucker, G.R., E. Hamayan and F. Genesee 1976. Affective, cognitive, and social factors in second language acquisition. *Canadian Modern Language Review*, 32, 214-226.

Uusitalo, S. 1987. *Zur Automatisierung der festen Wortverbindungen und zum Transfer bei den finnischen Deutschlernenden.* An unpublished Pro Gradu thesis. University of Jyväskylä.

VanPatten, B. 1989. Can learners attend to form and content while processing input? *Hispania*, 72, 409-417.

VanPatten, B. 1994. Evaluating the role of consciousness in second language acquisition: terms, linguistic features and research methodology. *AILA Review*, 11, 27-36.

Wardhaugh, R. 1970. The contrastive analysis hypothesis. *TESOL Quartely*, 4, 123-130.

White, L. 1991. Second language competence vs. second language performance: UG or processing strategies. In L. Eubank (ed.), *Point Counterpoint. Universal Grammar in the second language*. Amsterdam: John Benjamins, 167-197.

Widdowson, H.G. 1989 Knowledge of language and ability for use. *Applied Linguistics*, 10, 128-137.

Widiger, T.A., R.M. Knudson and L.G. Rorer 1980. Convergent and discriminant validity of measures of cognitive styles and abilities. *Journal of Personality and Social Psychology*, 39, 116-129.

Winitz, H. 1996 Grammaticality judgment as a function of explicit and implicit instruction in Spanish. *Modern Language Journal*, 80, 32-46.

Witkin, H.A., P.K. Oltman, E. Raskin and S.A. Karp 1971. *A manual for the embedded figures test.* Palo Alto, CA: Consulting Psychologists Press.

Witkin, H.A. and D.R. Goodenough and P.K. Oltman 1979. Psychological differentiation: current status. *Journal of Personality and Social Psychology*, 37, 1127-1145.

Witkin, H.A., C.A. Moore, D.R. Goodenaugh, P.W. Cox 1977a. Field-dependent and field-independent cognitive styles and their educational implications. *Review of Educational Research*, 47, 1-64.

Witkin, H.A., C.A. Moore, P.K. Oltman, D.R. Goodenaugh, F. Friedman, D.R. Owen, and E. Raskin 1977b. *Journal of Educational Psychology*, 69, 197-211.

Wode, H. 1978. Developmental sequencies in naturalistic L2 acquisition. In E. Hatch (ed.), *Second language acquisition.* Rowley, Mass.: Newbury House, 101-117.

Wode, H. 1981. *Learning a second language: an integrated view of language acquisition.* Heidelberg: Julius Groos.

Wulfeck, B. 1988. Grammaticality judgments and sentence comprehension in agrammatic aphasia. *Journal of Speech and Hearing Research*, 31, 72-81.

Wulfeck, B. 1993. A reaction time study of grammaticality judgments in children. *Journal of Hearing and Speech*, 36, 1208-1215.

Wulfeck, B. and E. Bates 1991. Differential sensitivity to errors of agreement and word order in Broca's aphasia. *Journal of Cognitive Neuroscience*, 3, 258-272.

Wulfeck, B. , E. Bates and R. Capasso 1991. A crosslinguistic study of grammaticality judgments in Broca's aphasia. Special Issue: Crosslinguistic studies of aphasia. Brain and Language, 4, 311-336.

Zobl, H. 1980. Develpmental and transfer errors; their common bases and (possibly) differential effects on subsequent learning. *TESOL Quartely*, 14, 469-479.

Zwitserlood, C.M.E. 1989. *Words and sentences: the effects of sentential-semantic context on spoken-word processing.* Unpublished doctoral dissertation. Nijmegen: The Catholic University of Nijmegen.

Appendix 1

EXPERIMENT 1

LIST OF TEST SENTENCES AND SUBJECTS' RESPONSES

TABLE 1 The mean RTs and judgments of the Finnish students of English to the English sentences (N = 20).

	Mean	SD	Judgment (Correct/ Incorrect)
101 It depends on him	2385	738	20/0
102 *The price depends of it	3125	1443	5/15
103 I visited Rome	2088	573	20/0
104 *We'll visit in London	2940	856	1/19
105 He has left for Bolton	2835	848	17/3
106 *My friend left to Wales	3160	1120	4/16
107 I found it in the fridge	2259	434	20/0
108 *Jim found it from the shop	3101	657	4/16
109 John likes her novels	2961	1381	18/2
110 *I like of biscuits	2278	1012	0/20
111 They fought hunger and poverty	2872	682	10/10
112 They fought against poverty	2780	586	19/1
113 We have just arrived at York	3129	1443	5/15
114 *The train arrived to London	2313	446	1/19
115 She left it in her car	2432	350	17/0
116 *He left it to Jill's room	3197	866	2/15
117 Jack never travels by train	2438	428	19/1
118 *We always travel with train	2819	686	8/10
119 We live in the country	2770	1738	18/2
120 *Dan lives at the country	2826	678	4/16
121 I lived in Paris for six years	3079	1203	18/2
122 John lived in Hull ten years	3769	2277	10/10
123 She lived in the 8th century	2631	553	17/0
124 *He lived on the 12th century	3263	919	5/12
125 Are you interested in cars?	2308	534	20/0
126 *Is she interested of cooking?	2918	653	2/18
127 What are you doing now?	2339	534	19/1
128 *What you are going to do now?	2867	677	10/10
129 They usually dine at home	2377	507	20/0
130 *We play usually baseball	3133	1198	5/15
131 Lil writes a letter every day	2720	747	20/0
132 *Bob reads every day a book	3454	1548	5/15
133 I gave him my photos yesterday	2879	860	20/0
134 *Hal gave me yesterday his bike	3559	972	12/8
135 Did she really become a nurse?	2809	556	20/0
136 *He wanted to become doctor	3019	775	5/15
137 Does she play the clarinet?	3073	926	19/1
138 *I play a violin	2471	1029	2/18
139 Richard loves wild nature	2596	753	19/1
140 *They pollute the nature	2796	548	15/5

141 She loves beautiful Rome	3057	1840	12/8
142 *He lived in the hot Morocco	3951	1777	13/7
143 Fruit is good for your health	3298	1039	17/3
144 *Fruits are expensive	3333	1653	15/5
145 Mary has good basic knowledge	2976	680	17/3
146 *His basic knowledges are good	3221	1213	3/17
147 Maggie had understood it	2721	430	17/2
148 *I have never understand it	2602	591	3/17
149 I have never swam before	3709	1305	13/6
150 Edward had not swum	3351	1218	14/6
151 *Ted drink tea every morning	2838	584	0/20
152 *My father have two brothers	2596	497	0/20
153 I was born in Leeds	2082	428	20/0
154 *Frank is born in 1955	3014	1806	6/14
155 We'll go to Derby on Friday	2688	441	20/0
156 We have a test next week	2974	1255	16/4
157 I have lived there since 1966	3056	693	17/3
158 *Ben lives in Cork since 1981	3207	978	9/11
159 She often goes to Italy	2491	696	19/1
160 We are often playing football	3395	1747	11/9
161 He wrote the play in 1980	2809	656	20/0
162 *He has sold his car in 1979	3347	1173	10/10
163 I heard her whistle	2451	1458	20/0
164 *He heard Jeff to come in	3303	1218	7/13
165 He never tries to stop me	2594	484	20/0
166 *We only try help them	2827	1063	2/18
167 I made him drink it	2487	675	20/0
168 *George made her to steal it	3373	1130	11/9
169 Sometimes we go dancing	2602	637	19/1
170 *We often go to fish	3221	1282	4/16
171 *We'll go swim tomorrow	2616	664	4/15
172 He was given a book	2066	564	20/0
173 *Him was given a letter	2798	1081	5/15
174 Bill is as clever as Jane	2263	338	20/0
175 *I am so tall as Ben	2510	620	1/19
176 The books sell easily	3053	973	15/5
177 *This car sells itself easily	2970	931	14/6
178 She makes good coffee	2303	600	18/2
179 *Larry cooks good coffee	2761	676	5/15
180 It will be the death of you	3727	1425	11/9
181 *You will be the death for me	3621	1173	7/13
182 Ricky made fun of me	2353	649	20/0
183 *They did fun of him	2750	744	7/13
184 *My finnish friends are coming	3063	713	8/12
185 He bought Finnish glass	2815	586	20/0
186 *He gave me his address	2692	705	14/6
187 What is your address?	2225	647	19/0
188 We saw a very good film today	3585	889	16/4
189 Mrs Darryll loves her children	2393	408	20/0
190 *David didn't see I on Monday	2729	338	2/18
191 *Her is two department stores	3234	450	0/20

TABLE 2 The mean RTs and judgments of the native speakers of English to the English sentences (N = 15).

	Mean	SD	Judgment (Correct/ Incorrect)
101 It depends on him	1617	447	15/0
102 *The price depends of it	1780	527	1/14
103 I visited Rome	1298	391	15/0
104 *We'll visit in London	1991	427	3/12
105 He has left for Bolton	1983	362	15/0
106 *My friend left to Wales	1921	470	1/14
107 I found it in the fridge	1805	422	15/0
108 *Jim found it from the shop	2491	610	2/13
109 John likes her novels	2014	635	14/1
110 *I like of biscuits	1285	309	0/15
111 They fought hunger and poverty	2589	1586	15/0
112 *They fought against poverty	1741	505	15/0
113 We have just arrived at York	2425	648	8/7
114 *The train arrived to London	1637	308	0/15
115 She left it in her car	1986	615	15/0
116 *He left it to Jill's room	2034	428	0/15
117 Jack never travels by train	1827	367	15/0
118 *We always travel with train	2148	1081	2/13
119 We live in the country	1763	646	15/0
120 *Dan lives at the country	1760	460	2/13
121 I lived in Paris for six years	1745	427	15/0
122 John lived in Hull ten years	2407	496	11/4
123 She lived in the 8th century	2019	629	15/0
124 *He lived on the 12th century	2111	321	0/15
125 Are you interested in cars?	1616	346	15/0
126 *Is she interested of cooking?	1861	397	0/15
127 What are you doing now?	1571	311	15/0
128 *What you are going to do now?	1844	501	5/10
129 They usually dine at home	1584	284	15/0
130 *We play usually baseball	1837	452	0/15
131 Lil writes a letter every day	1898	391	15/0
132 *Bob reads every day a book	2387	692	3/12
133 I gave him my photos yesterday	2160	502	15/0
134 *Hal gave me yesterday his bike	2282	454	2/13
135 Did she really become a nurse?	1831	422	15/0
136 *He wanted to become doctor	1673	528	5/10
137 Does she play the clarinet?	1827	363	15/0
138 *I play a violin	2567	1283	7/8
139 Richard loves wild nature	2424	933	10/4
140 *They pollute the nature	1996	448	5/10
141 She loves beautiful Rome	2609	977	10/5
142 *He lived in the hot Morocco	2257	761	3/12
143 Fruit is good for your health	2083	569	14/1
144 *Fruits are expensive	2038	692	9/6
145 Mary has good basic knowledge	2621	955	13/2
146 *His basic knowledges are good	1897	451	0/15
147 Maggie had understood it	2473	890	13/2

148 *I have never understand it	2006	496	3/12
149 *I have never swam before	2460	669	7/8
150 Edward had not swum	3005	1857	9/6
151 *Ted drink tea every morning	2050	599	0/12
152 *My father have two brothers	1522	238	1/14
153 I was born in Leeds	1317	229	15/0
154 *Frank is born in 1955	1898	887	1/14
155 We'll go to Derby on Friday	1954	388	14/1
156 We have a test next week	1654	487	13/2
157 I have lived there since 1966	2098	501	14/1
158 *Ben lives in Cork since 1981	2203	478	1/14
159 She often goes to Italy	1604	378	14/1
160 We are often playing football	2234	676	9/6
161 He wrote the play in 1980	1814	633	15/0
162 *He has sold his car in 1979	2606	1209	3/12
163 I heard her whistle	1430	334	15/0
164 *He heard Jeff to come in	1905	454	0/15
165 He never tries to stop me	1836	307	15/0
166 *We only try help them	2074	744	3/12
167 I made him drink it	1635	311	14/1
168 *George made her to steal it	2060	609	0/15
169 Sometimes we go dancing	1758	374	15/0
170 *We often go to fish	2201	867	6/9
171 *We'll go swim tomorrow	2391	1038	7/8
172 He was given a book	1486	327	14/1
173 *Him was given a letter	1693	490	2/13
174 Bill is as clever as Jane	1848	846	14/1
175 *I am so tall as Ben	1817	331	1/14
176 The books sell easily	2205	825	15/0
177 *This car sells itself easily	2643	525	8/7
178 She makes good coffee	1475	421	15/0
179 *Larry cooks good coffee	2266	1246	6/9
180 It will be the death of you	2478	669	14/1
181 *You will be the death for me	2524	1079	1/14
182 Ricky made fun of me	1870	588	15/0
183 *They did fun of him	1544	312	1/14
184 *My finnish friends are coming	2172	535	9/6
185 He bought Finnish glass	2168	878	14/1
186 *He gave me his address	2661	971	8/7
187 What is your address?	1525	288	15/0
188 We saw a very good film today	2169	512	14/1
189 Mrs Darryll loves her children	1781	581	15/0
190 *David didn't see I on Monday	1999	486	1/14
191 *Her is two department stores	2103	608	0/15

TABLE 3 The mean RTs and judgments of the German teachers to the English sentences (N=13).

	Mean	SD	Judgment (Correct/ Incorrect)
101 It depends on him	2383	491	12/1
102 *The price depends of it	3028	883	2/11
103 I visited Rome	2081	442	12/1
104 *We'll visit in London	3697	2163	0/13
105 He has left for Bolton	3887	2842	8/5
106 *My friend left to Wales	3512	1399	7/6
107 I found it in the fridge	2845	780	13/0
108 *Jim found it from the shop	4251	2315	3/10
109 John likes her novels	3483	2084	11/2
110 *I like of biscuits	2443	780	0/13
111 They fought hunger and poverty	3360	1749	4/9
112 They fought against poverty	2514	528	13/0
113 We have just arrived at York	3904	1797	9/4
114 *The train arrived to London	3093	2160	3/10
115 She left it in her car	3482	2223	11/2
116 *He left it to Jill's room	3985	1492	3/10
117 Jack never travels by train	3186	1640	12/1
118 *We always travel with train	3241	974	2/11
119 We live in the country	3175	1039	1/12
120 *Dan lives at the country	2864	727	1/12
121 I lived in Paris for six years	2843	718	11/2
122 John lived in Hull ten years	3434	1027	6/7
123 She lived in the 8th century	2831	651	13/0
124 *He lived on the 12th century	2995	648	0/13
125 Are you interested in cars?	2081	442	12/1
126 *Is she interested of cooking?	2890	718	1/12
127 What are you doing now?	2565	438	13/0
128 *What you are going to do now?	3712	1868	5/8
129 They usually dine at home	3239	915	10/3
130 *We play usually baseball	3984	2801	9/4
131 Lil writes a letter every day	3028	759	12/1
132 *Bob reads every day a book	3849	2099	6/7
133 I gave him my photos yesterday	3213	1436	13/0
134 *Hal gave me yesterday his bike	3578	1007	5/8
135 Did she really become a nurse?	3132	910	13/0
136 *He wanted to become doctor	3083	1533	2/11
137 Does she play the clarinet?	3564	1440	13/0
138 *I play a violin	3954	2798	3/10
139 Richard loves wild nature	3268	919	12/1
140 *They pollute the nature	3742	1315	11/2
141 She loves beautiful Rome	4301	2303	7/5
142 *He lived in the hot Morocco	4141	2121	7/6
143 Fruit is good for your health	4251	2261	9/4
144 *Fruits are expensive	2627	1421	11/2
145 Mary has good basic knowledge	4751	2807	11/2

146 *His basic knowledges are good	4422	2415	9/4
147 Maggie had understood it	4568	2199	10/3
148 *I have never understand it	3409	1200	4/9
149 *I have never swam before	4764	2354	7/6
150 Edward had not swum	5134	3021	5/8
151 *Ted drink tea every morning	2964	974	1/12
152 *My father have two brothers	2467	642	0/13
153 I was born in Leeds	2873	2105	13/0
154 *Frank is born in 1955	3351	2125	7/6
155 We'll go to Derby on Friday	3237	1088	10/3
156 We have a test next week	2914	852	12/1
157 I have lived there since 1966	3425	944	10/3
158 *Ben lives in Cork since 1981	3532	1329	10/3
159 She often goes to Italy	2476	598	13/0
160 We are often playing football	3292	2124	11/1
161 He wrote the play in 1980	2731	979	12/1
162 *He has sold his car in 1979	3480	1145	11/2
163 I heard her whistle	3792	2847	9/4
164 *He heard Jeff to come in	4028	1642	5/8
165 He never tries to stop me	2782	421	12/1
166 *We only try help them	2455	589	1/12
167 I made him drink it	3762	2477	7/6
168 *George made her to steal it	3835	1559	7/6
169 Sometimes we go dancing	3283	1778	11/2
170 *We often go to fish	3898	2670	4/9
171 *We'll go swim tomorrow	3673	1762	3/10
172 He was given a book	3832	2928	8/5
173 *Him was given a letter	3279	1554	0/13
174 Bill is as clever as Jane	2524	683	12/1
175 *I am so tall as Ben	3553	1613	5/8
176 The books sell easily	4044	1409	5/8
177 *This car sells itself easily	4034	2380	5/8
178 She makes good cofee	3256	2151	10/3
179 *Larry cooks good coffee	3669	2300	12/1
180 It will be the death of you	4587	2563	5/8
181 *You will be the death for me	4408	1732	5/8
182 Ricky made fun of me	3387	2160	10/3
183 *They did fun of him	4023	2534	1/12
184 *My finnish friends are coming	2819	762	13/0
185 He bought Finnish glass	3885	1841	12/1
186 *He gave me his address	3004	658	12/1
187 What is your address?	2647	598	12/1
188 We saw a very good film today	3239	885	9/4
189 Mrs Darryll loves her children	2727	740	12/1
190 *David didn't see I on Monday	3102	791	0/13
191 *Her is two department stores	3134	700	1/12

TABLE 4 The mean RTs and judgments of the German students of English to the English sentences (N=20).

	Mean	SD	Judgments (Correct/ Incorrect)
101 It depends on him	1812	661	17/3
102 *The price depends of it	2399	1498	7/13
103 I visited Rome	1701	496	18/2
104 *We'll visit in London	2334	810	5/15
105 He has left for Bolton	2527	1314	11/9
106 *My friend left to Wales	2310	869	9/11
107 I found it in the fridge	1942	664	20/0
108 *Jim found it from the shop	2045	441	3/17
109 John likes her novels	2443	1923	16/4
110 *I like of biscuits	1574	304	0/20
111 They fought hunger and poverty	2395	1022	15/5
112 They fought against poverty	1910	581	18/2
113 We have just arrived at York	2225	1513	19/1
114 *The train arrived to London	1857	379	2/18
115 She left it in her car	2300	953	18/2
116 *He left it to Jill's room	2510	935	7/13
117 Jack never travels by train	2051	836	18/2
118 *We always travel with train	2096	876	6/13
119 We live in the country	2054	1963	16/4
120 *Dan lives at the country	2069	624	7/13
121 I lived in Paris for six years	2046	438	12/8
122 John lived in Hull ten years	2464	729	6/13
123 She lived in the 8th century	2561	1087	18/2
124 *He lived on the 12th century	2273	742	3/17
125 Are you interested in cars?	1708	383	18/2
126 *Is she interested of cooking?	2317	804	4/15
127 What are you doing now?	1885	664	18/2
128 *What you are going to do now?	2352	736	8/12
129 They usually dine at home	1845	583	17/3
130 *We play usually baseball	2483	1185	10/10
131 Lil writes a letter every day	1860	368	20/0
132 *Bob reads every day a book	2391	880	6/14
133 I gave him my photos yesterday	2247	812	20/0
134 *Hal gave me yesterday his bike	2297	629	9/10
135 Did she really become a nurse?	2570	1197	18/2
136 *He wanted to become doctor	2126	608	9/11
137 Does she play the clarinet?	2154	735	20/0
138 *I play a violin	2236	910	8/12
139 Richard loves wild nature	2011	627	19/1
140 *They pollute the nature	2087	739	16/4
141 She loves beautiful Rome	2475	1924	14/6
142 *He lived in the hot Morocco	2497	811	14/5
143 Fruit is good for your health	1955	443	18/2
144 *Fruits are expensive	2207	1704	18/2
145 Mary has good basic knowledge	1966	576	18/1
146 *His basic knowledges are good	2279	509	13/7

147 Maggie had understood it	2858	1932	11/9
148 *I have never understand it	2105	917	4/15
149 *I have never swam before	2326	735	11/9
150 Edward had not swum	2552	1010	7/13
151 *Ted drink tea every morning	2177	743	10/8
152 *My father have two brothers	1874	452	5/15
153 I was born in Leeds	1605	500	20/0
154 *Frank is born in 1955	2351	1239	11/9
155 We'll go to Derby on Friday	1811	447	20/0
156 We have a test next week	2314	968	12/7
157 I have lived there since 1966	2220	714	14/5
158 *Ben lives in Cork since 1981	2257	745	10/10
159 She often goes to Italy	1810	738	20/0
160 We are often playing football	2041	697	17/3
161 He wrote the play in 1980	2092	624	18/2
162 *He has sold his car in 1979	2746	1900	11/9
163 I heard her whistle	1882	772	17/3
164 *He heard Jeff to come in	2330	1028	6/14
165 He never tries to stop me	2030	847	17/3
166 *We only try help them	2148	943	3/17
167 I made him drink it	2189	1201	12/8
168 *George made her to steal it	2330	539	10/10
169 Sometimes we go dancing	2076	946	19/1
170 *We often go to fish	2141	1201	4/16
171 *We'll go swim tomorrow	2091	766	12/8
172 He was given a book	1885	752	18/2
173 *Him was given a letter	2183	1133	6/14
174 Bill is as clever as Jane	1796	402	18/2
175 *I am so tall as Ben	2572	1792	3/17
176 The books sell easily	2122	631	9/11
177 *This car sells itself easily	2655	1443	10/10
178 She makes good coffee	1975	618	18/2
179 *Larry cooks good coffee	2091	780	14/6
180 It will be the death of you	2487	823	11/9
181 *You will be the death for me	2516	870	9/10
182 Ricky made fun of me	1966	704	15/5
183 *They did fun of him	2274	859	3/17
184 *My finnish friends are coming	2666	920	15/4
185 He bought Finnish glass	2460	981	15/5
186 *He gave me his address	1930	905	17/3
187 What is your address?	1645	479	20/0
188 We saw a very good film today	2485	1286	16/4
189 Mrs Darryll loves her children	1807	387	19/1
190 *David didn't see I on Monday	2329	512	5/15
191 *Her is two department stores	2093	688	4/13

TABLE 5 The mean RTs and judgments of the Finnish students of English to the German sentences (N = 20).

	Mean	SD	Judgment (Correct/ Incorrect)
501 Ute wohnte drei Jahre in Köln	3719	1026	15/5
502 Er wohnte für ein Jahr in Ulm	4019	1195	13/7
503 Er fand es in einem Geschäft	3008	947	18/2
504 *Lisa fand es von einem Schrank	3485	1026	3/17
505 Eva ist nach Bonn abgefahren	3499	978	18/2
506 *Franz hat für Kiel verlassen	3639	1301	8/12
507 Sie kam in Bremen an	3265	1484	20/0
508 *Hans kam nach Leipzig an	3463	1287	14/6
509 Ich war interessiert an Tieren	3797	1461	9/11
510 *Er war interessiert in Katzen	3536	1102	12/8
511 Mir wurde ein Auto gegeben	3455	976	10/10
512 *Ich wurde ein Buch gegeben	3806	1108	11/9
513 Franz wurde Polizist	2746	1035	14/6
514 *Petra wurde eine Verkäuferin	3764	1663	12/8
515 Sie wohnt in Basel seit 1930	3091	965	18/2
516 Er hat in Ulm seit 1951 gewohnt	3830	1588	14/6
517 Karl hörte uns spielen	2588	784	16/4
518 *Er hörte Ellen zu singen	3573	1853	8/12
519 Günther sah Olga weinen	2994	975	19/1
520 *Lola sah ihn zu gehen	3081	1215	7/13
521 Ilse stürzte ins Zimmer	3241	931	18/2
522 *Er raschte zur Bank	3063	1175	16/4
523 Uwe macht guten Kaffee	2876	962	16/4
524 Linda kocht starken Kaffee	3318	1084	15/5
525 Karl ist pünktlicher als du	3368	1238	18/2
526 *Elsa iste mehr tapfer als ich	3942	1850	15/5
527 Willi versteht sie nicht	3449	1363	16/4
528 *Uwe verstandet es nicht	3597	1119	8/12
529 Ich bin in Kiel geboren	2602	827	19/1
530 *Ich war in Bonn geboren	3726	1342	9/11
531 *Zwei Männer har Wörterbücher	3687	1374	5/15
532 *Ich hast heute dort gesitzen	3272	913	5/15
533 Herr Müller hat zwei Bücher	3023	1247	19/1
534 Mein Name ist Inga Bergmann	3350	1310	17/3

TABLE 6 The mean RTs and judgments of the German teachers to the German sentences (N = 13).

	Mean	SD	Judgment (Correct/ Incorrect)
501 Ute wohnte drei Jahre in Köln	2987	2177	13/0
502 Er wohnte für ein Jahr in Ulm	3732	2446	8/5
503 Er fand es in einem Geschäft	2315	420	13/0
504 *Lisa fand es von einem Schrank	2799	764	0/13
505 Eva ist nach Bonn abgefahren	2873	665	12/1
506 *Franz hat fur Kiel verlassen	2238	576	0/13
507 Sie kam in Bremen an	1873	371	13/0
508 *Hans kam nach Leipzig an	2005	387	0/13
509 Ich war interessiert an Tieren	3020	1139	11/2
510 *Er war interessiert in Katzen	2359	798	1/12
511 Mir wurde ein Auto gegeben	4097	2611	10/3
512 *Ich wurde ein Buch gegeben	2237	484	0/13
513 Franz wurde Polizist	1697	478	13/0
514 *Petra wurde eine Verkäuferin	2902	820	7/6
515 Sie wohnt in Basel seit 1930	3520	2185	9/4
516 Er hat in Ulm seit 1951 gewohnt	3415	1320	4/8
517 Karl hörte uns spielen	1865	420	13/0
518 *Er hörte Ellen zu singen	2275	499	0/13
519 Günther sah Olga weinen	2025	482	12/0
520 *Lola sah ihn zu gehen	2095	344	1/12
521 Ilse stürzte ins Zimmer	1822	470	13/0
522 *Er raschte zur Bank	1765	417	0/13
523 Uwe macht guten Kaffee	1927	429	13/0
524 Linda kocht starken Kaffee	2085	637	13/0
525 Karl ist pünktlicher als du	2295	448	13/0
526 *Elsa iste mehr tapfer als ich	2020	393	0/13
527 Willi versteht sie nicht	2017	624	13/0
528 *Uwe verstandet es nicht	1759	494	1/12
529 Ich bin in Kiel geboren	2154	810	12/0
530 *Ich war in Bonn geboren	2414	798	1/12
531 *Zwei Männer har Wörterbücher	2039	543	0/13
532 *Ich hast heute dort gesitzen	2009	570	0/13
533 Herr Müller hat zwei Bücher	2421	779	13/0
534 Mein Name ist Inga Bergmann	2002	676	13/0

TABLE 7 The mean RTs and judgments of the German students of English to the German sentences (N = 20).

	Mean	SD	Judgments (Correct/ Incorrect)
501 Ute wohnte drei Jahre in Köln	1604	347	18/2
502 Er wohnte für ein Jahr in Ulm	1803	680	17/3
503 Er fand es in einem Geschäft	1730	508	18/2
504 *Lisa fand es von einem Schrank	2150	621	5/14
505 Eva ist nach Bonn abgefahren	1937	1049	16/4
506 *Franz hat für Kiel verlassen	1821	665	3/17
507 Sie kam in Bremen an	1402	286	20/0
508 *Hans kam nach Leipzig an	1855	342	2/18
509 Ich war interessiert an Tieren	2523	1415	15/5
510 *Er war interessiert in Katzen	2239	949	4/16
511 Mir wurde ein Auto gegeben	2109	885	17/3
512 *Ich wurde ein Buch gegeben	1832	553	0/20
513 Franz wurde Polizist	1160	279	19/1
514 *Petra wurde eine Verkäuferin	1912	687	15/5
515 Sie wohnt in Basel seit 1930	1910	704	20/0
516 Er hat in Ulm seit 1951 gewohnt	2684	1662	14/6
517 Karl hörte uns spielen	1825	811	17/3
518 *Er hörte Ellen zu singen	1873	613	1/19
519 Günther sah Olga weinen	1468	327	20/0
520 *Lola sah ihn zu gehen	1931	486	1/19
521 Ilse stürzte ins Zimmer	1327	273	20/0
522 *Er raschte zur Bank	1674	388	1/19
523 Uwe macht guten Kaffee	1672	846	19/1
524 Linda kocht starken Kaffee	1533	361	20/0
525 Karl ist pünktlicher als du	1635	365	18/2
526 *Elsa ist mehr tapfer als ich	1930	441	1/19
527 Willi versteht sie nicht	1576	597	19/1
528 *Uwe verstandet es nicht	1598	304	6/14
529 Ich bin in Kiel geboren	1427	372	20/0
530 *Ich war in Bonn geboren	1999	818	4/16
531 *Zwei Männer hat Wörterbücher	1550	319	4/16
532 *Ich hast heute dort gesitzen	2166	1019	1/19
533 Herr Müller hat zwei Bücher	1526	388	19/1
534 Mein Name ist Inga Bergmann	1279	282	19/0

TABLE 8 The mean RTs and judgments of the Finnish students of English to the English/ German translation pairs (N = 20).

	Mean	SD	Judgment (Correct/ Incorrect)
401 I lived here for five years Ich wohnte hier für fünf Jahre	4476	1764	11/9
402 He found it in my room *Er fand es aus meinem Zimmer	4582	2236	6/13
403 I have left for Paris *Ich habe für Paris verlassen	3791	1239	6/13
404 The bus arrived at Glasgow *Der Bus kam nach Glasgow an	3825	976	10/10
405 I came to Yuma yesterday *Ich kam gestern nach Yuma an	4043	1155	12/8
406 Eve was interested in cars *Eve war interessiert in Autos	4140	970	11/9
407 He was given a book *Er wurde ein Buch gegeben	3604	1067	12/8
408 She became a teacher *Sie wurde eine Lehrerin	4210	1838	12/7
409 He has lived here since 1980 *Er hat hier seit 1980 gewohnt	4025	934	14/6
410 Jim heard him speak *Jim hörte ihn zu sprechen	3891	1007	8/12
411 Helen saw me come Helen sah mich kommen	3918	1247	14/6
412 They rushed to the station *Sie raschten zum Bahnhof	3685	938	17/3
413 She makes good coffee Sie macht guten Kaffee	4446	1921	14/6
414 She is more punctual than you *Sie ist mehr pünktlich als du	4548	1618	17/3
415 I can't understand him *Ich kann ihn nicht verstanden	3386	742	16/4
416 Doctor Berger opened the door Doktor Berger öffnete die Tür	3757	1123	16/4
417 She lived in the 11th century Sie lebte im 11. Jahrhundert	3694	735	16/3
418 I have already seen Berlin Ich habe schon Berlin gesehen	3522	867	18/2
419 She has visited Berlin Sie hat Berlin besucht	3563	1247	18/2
420 The boy ran across the street *Der Junge lief auf der Strasse	4038	1177	3/17
421 Mrs Wolters closed the window *Frau Wolters öffnete die Tür	3598	879	1/19
422 On Friday I'll be in Ankara Am Freitag bin ich in Ankara	4314	1053	14/6
423 I was born in Wales *Ich war in Wales geboren	4142	1292	13/6

TABLE 9 The mean RTs and judgments of the German students to the English/ German translation pairs (N=13).

	Mean	SD	Judgment (Correct/ Incorrect)
401 I lived here for five years Ich wohnte hier für fünf Jahre	5444	2299	5/7
402 He found it in my room *Er fand es aus meinem Zimmer	4083	2100	0/13
403 I have left for Paris *Ich habe für Paris verlassen	2726	785	0/13
404 The bus arrived at Glasgow *Der Bus kam nach Glasgow an	2754	795	0/13
405 I came to Yuma yesterday *Ich kam gestern nach Yuma an	3514	1768	4/9
406 Eve was interested in cars *Eve war interessiert in Autos	3863	1395	2/11
407 He was given a book *Er wurde ein Buch gegeben	2481	619	1/12
408 She became a teacher *Sie wurde eine Lehrerin	4777	2504	9/3
409 He has lived here since 1980 *Er hat hier seit 1980 gewohnt	4024	2200	9/4
410 Jim heard him speak *Jim hörte ihn zu sprechen	3403	859	0/13
411 Helen saw me come Helen sah mich kommen	2848	645	11/2
412 They rushed to the station *Sie raschten zum Bahnhof	3023	660	1/12
413 She makes good coffee Sie macht guten Kaffee	2703	491	12/1
414 She is more punctual than you *Sie ist mehr pünktlich als du	3653	1378	2/11
415 I can't understand him *Ich kann ihn nicht verstanden	2985	590	3/10
416 Doctor Berger opened the door Doktor Berger öffnete die Tür	2801	646	11/2
417 She lived in the 11th century Sie lebte im 11. Jahrhundert	3002	545	10/3
418 I have already seen Berlin Ich habe schon Berlin gesehen	3766	2152	6/6
419 She has visited Berlin Sie hat Berlin besucht	3313	959	13/0
420 The boy ran across the street *Der Junge lief auf der Strasse	4468	1679	2/11
421 Mrs Wolters closed the window *Frau Wolters öffnete die Tür	3333	1374	4/9
422 On Friday I'll be in Ankara Am Freitag bin ich in Ankara	3596	1064	10/3
423 I was born in Wales *Ich war in Wales geboren	3540	1529	4/9

TABLE 10 The mean RTs and judgments of the German students to the English/ German translation pairs (N = 20).

	Mean	SD	Judgment (Correct/ Incorrect)
401 I lived here for five years Ich wohnte hier für fünf Jahre	2411	900	14/6
402 He found it in my room *Er fand es aus meinem Zimmer	3145	1545	14/5
403 I have left for Paris *Ich habe für Paris verlassen	2180	1004	6/14
404 The bus arrived at Glasgow *Der Bus kam nach Glasgow an	2166	675	10/10
405 I came to Yuma yesterday *Ich kam gestern nach Yuma an	2398	1019	12/8
406 Eve was interested in cars *Eve war interessiert in Autos	2849	1803	13/7
407 He was given a book *Er wurde ein Buch gegeben	2225	715	4/16
408 She became a teacher *Sie wurde eine Lehrerin	2052	768	16/3
409 He has lived here since 1980 *Er hat hier seit 1980 gewohnt	2133	631	13/7
410 Jim heard him speak *Jim hörte ihn zu sprechen	2434	895	4/16
411 Helen saw me come Helen sah mich kommen	2616	1968	15/5
412 They rushed to the station *Sie raschten zum Bahnhof	2513	1070	7/13
413 She makes good coffee Sie macht guten Kaffee	1753	404	19/1
414 She is more punctual than you *Sie ist mehr pünktlich als du	2513	978	7/13
415 I can't understand him *Ich kann ihn nicht verstanden	2282	1113	10/9
416 Doctor Berger opened the door Doktor Berger öffnete die Tür	1921	859	19/1
417 She lived in the 11th century Sie lebte im 11. Jahrhundert	2025	509	19/1
418 I have already seen Berlin Ich habe schon Berlin gesehen	2526	1034	18/2
419 She has visited Berlin Sie hat Berlin besucht	2007	506	18/2
420 The boy ran across the street *Der Junge lief auf der Strasse	2308	894	13/7
421 Mrs Wolters closed the window *Frau Wolters öffnete die Tür	2310	819	6/14
422 On Friday I'll be in Ankara Am Freitag bin ich in Ankara	2155	742	17/3
423 I was born in Wales *Ich war in Wales geboren	2112	762	14/6

TABLE 11 The mean RTs and judgments of the Finnish students of English to the Swedish sentences (N=20).

	Mean	SD	Judgment (Correct/ Incorrect)
701 Jag bodde här för många år	3609	1457	13/7
702 Kaj bodde i Lund i många år	3452	1441	11/9
703 Vi hittade det i kylskåpet	2362	755	20/0
704 *Ulla hittade det från en låda	3696	1869	1/19
705 *Olle lämnade det till min bil	3027	430	2/18
706 Hon lämnade boken i din väska	2769	842	20/0
707 Hon levde på 1500-talet	2431	938	20/0
708 *Gustav III levde i 1700-talet	3350	1385	8/12
709 Per har begett sig till Jbo	3545	1242	19/1
710 Nils har åkt till Skåne	2829	1090	14/1
711 *Lasse har lämnat för Torneå	3497	1581	8/12
712 Lena är intresserad av hundar	2732	1038	17/3
713 *Han är intresserad i fåglar	2776	943	10/10
714 Jag är född i Umeå	2535	756	20/0
715 *Jag var född år 1951	3277	868	6/14
716 *Hon rasade till Stockmann	2785	819	17/3
717 Pelle rusade in i huset	2675	728	17/1
718 Pia kokar gott kaffe	2742	1049	18/2
719 *Lars makar starkt kaffe	3012	829	4/16
720 Hon såg mig läsa	2708	927	19/1
721 *Ulf såg dig att simma	3430	1052	7/13
722 *Jag har aldrig påstod detta	3845	1717	15/5
723 Hans har aldrig påstått detta	3687	1133	15/5
724 Kalle blev mekaniker	3032	1335	13/7
725 *Lotta blev en lärarinna	3314	2021	10/10
726 Gustafsson öppnade fönstret	2834	1280	19/1
727 Jag läser inte hennes böcker	2808	978	18/2
728 *Jag kommer tillbaka i går	2872	924	1/19
729 *Du har sett han många gånger	3318	698	3/17

TABLE 12 The mean RTs and judgments of the Finnish students of English to the English/Swedish translation pairs (N = 20).

	Mean	SD	Judgment (Correct/ Incorrect)
601 Carl lives in the country	3413	1182	19/1
Carl bor på landet			
602 We lived in Yale for two years	3820	773	16/4
*Vi bodde i Yale för två år			
603 We found it in his pocket	3354	747	18/1
Vi hittade det i hans ficka			
604 She left it in my room	3739	1127	3/17
*Hon lämnade det till mitt rum			
605 He lived in the 16th century	4076	1324	10/10
*Han levde i 1500-talet			
606 They have left for Dublin	4217	1347	8/12
*De har lämnat för Dublin			
607 Len is interested in films	3966	885	12/8
*Len är intresserad i filmer			
608 I was born in 1943	3873	1231	8/12
*Jag var född år 1943			
609 Lisa rushed to the park	3631	1008	18/2
*Lisa rasade till parken			
610 I make strong coffee	3366	632	7/13
*Jag makar starkt kaffe			
611 On Monday she'll be in Cork	4345	1404	14/6
På måndag är hon i Cork			
612 Ben heard us go out	3661	814	17/2
Ben hörde oss gå ut			
613 I saw her die	3623	1271	7/13
*Jag såg henne att dö			
614 I have never understood him	4089	1808	14/6
*Jag har aldrig förstod honom			
615 He became a doctor	3988	1579	10/10
*Han blev en läkare			
616 Nielsen has read many books	3376	691	20/0
Nielsen har läst många böcker			
617 They have bought two tickets	3288	635	20/0
De har köpt två biljetter			
618 Professor Jonson flew to Ohio	4282	1793	3/17
* Professor Jonson reste hem			
619 Eve cannot play the violin	3722	676	5/15
*Eve kan inte spela piano			

TABLE 13 The mean RTs and judgments of the Finnish students of English to the Finnish sentences (N = 20).

	Mean	SD	Judgment (Correct/ Incorrect)
301 ?Hän asuu Turussa vuodesta 56	2810	1023	5/14
302 ?Olen asunut täällä vuodesta 63	2153	619	20/0
303 ?Se kirja myy nopesti	2529	813	11/7
304 ?Tuo malli myy hyvin	2220	570	18/2
305 *En ottanut sille huomioon	2381	789	1/19
306 *Pekka ei mennyt eilen kotona	2300	421	0/20
307 Hän sulki olohuoneen ikkunan	2082	615	19/1
308 Keskiviikkona ostin kirjoja	2793	1666	19/1

TABLE 14 The mean RTs and judgments of the Finnish students of English to the English/Finnish translation pairs (N = 20).

	Mean	SD	Judgment (Correct/ Incorrect)
201 He lived in York for two years ?Hän eli Yorkissa kaksi vuotta	3634	943	17/3
202 Dan has lived here since 1972 ?Dan asuu täällä vuodesta 1972	3695	1002	7/13
203 The books sell quickly ?Ne kirjat myyvät nopeasti	3718	1137	13/7
204 Ann Philips bought two books Ann Philips osti kaksi kirjaa	3099	576	19/1
205 James wanted to come today James halusi tulla tänään	3406	958	19/1
206 She forgot her husband's keys *Hän unohti miehensä kellon	3286	704	0/20
207 Bill walked across the river *Bill käveli jokea pitkin	3287	789	3/16

Appendix 2

LIST OF TEST SENTENCES AND THE LEARNERS' RESPONSES

TABLE 1 Mean RTs and judgments of English sentences by the NSs of English (N=5).

Target Sentences	Mean	SD	Judgment (Acceptable/ Unacceptable)
102 *Amy had a golden ring in her finger	2311	618	0/5
104 *He saw his face at the mirror	2062	596	0/5
106 *She thought Jack was in the phone	2407	740	0/5
108 You have a new ring on your finger	1998	740	5/0
109 *Ann saw a hat on the shop window	2252	690	1/4
110 Jill noticed Bill in the mirror	2283	573	5/0
111 He heard them talking on the phone	1831	382	5/0
113 *Jim could see no handle in the door	2401	930	0/5
114 *Ted thought his hair was in fire	2282	739	0/5
115 She liked the shoes in the window	2440	974	5/0
117 *Both countries are now in war	2635	848	0/5
118 *They wanted to sit in her table	2073	707	0/5
119 There was no handle on the door	1537	472	4/1
121 France and Germany were at war	2015	885	5/0
123 She likes to sit at Joe's table	1895	469	5/0
124 They said our house was on fire	2151	680	5/0
Control Sentences			
101 William Davies was born in 1940	2037	480	5/0
103 *She lives in London since 1987	2384	1057	0/5
105 John was interested in cars	1677	544	5/0
107 *George Smith is born in 1942	2546	1787	1/4
112 He has lived in Paris since 1986	1961	562	5/0
116 *Jack left to Chicago last Friday	1974	720	0/5
120 Betty left for Boston on Monday	2171	1129	5/0
122 *Janice was interested of cooking	1717	534	0/5

TABLE 2 Mean RTs and judgments of English L2 sentences by the Finns (N=21).

Target Sentences	Mean	SD	Judgment (Acceptable/ Unacceptable)
102 *Amy had a golden ring in her finger	2841	554	19/2
104 *He saw his face at the mirror	2811	836	12/9
106 *She thought Jack was in the phone	2483	734	10/11
108 You have a new ring on your finger	2869	802	14/7
109 *Ann saw a hat on the shop window	3317	1201	13/8
110 Jill noticed Bill in the mirror	2988	1199	16/4
111 He heard them talking on the phone	2375	563	18/3
113 *Jim could see no handle in the door	3187	976	16/5
114 *Ted thought his hair was in fire	3090	884	15/6
115 She liked the shoes in the window	2985	757	18/3
117 *Both countries are now in war	2965	1000	17/3
118 *They wanted to sit in her table	3177	1032	8/12
119 There was no handle on the door	2894	1327	12/8
121 France and Germany were at war	2493	476	18/3
123 She likes to sit at Joe's table	2898	1181	14/6
124 They said our house was on fire	2519	866	20/1
Control sentences			
101 William Davies was born in 1940	2016	382	21/0
103 *She lives in London since 1987	2368	1005	5/15
105 John was interested in cars	2081	513	21/0
107 *George Smith is born in 1942	2130	464	2/19
112 He has lived in Paris since 1986	2404	585	19/2
116 *Jack left to Chicago last Friday	2715	676	11/10
120 Betty left for Boston on Monday	2341	831	17/4
122 *Janice was interested of cooking	2092	526	3/18

TABLE 3 Mean RTs and judgments of Swedish L3 sentences by the Finns (N=21).

Target Sentences	Mean	SD	Judgment Acceptable/ Unacceptable
201 *Då bodde Kalle om landet	2889	593	5/16
202 *Lena talade ivrigt på telefon	3186	1036	15/6
204 *Pelle drog henne på håret	2937	1142	13/8
205 *De bodde i ett fint hotell	3126	1203	15/4
206 *Just nu bor de på utlandet	3498	1214	14/6
208 Nu bor Stina på landet	2196	754	20/1
210 Birgitta talar alltid i telefon	2955	1364	14/5
211 *Hon hade inga skor i fötterna	2979	806	4/17
212 Pojken drog Lisa i håret	2702	686	11/10
213 Lars bodde på ett nytt hotell	2448	895	15/6
214 I somras var hon i utlandet	3211	1152	13/8
215 *Per mötte henne i tåget till Åbo	3465	1278	13/7
217 Stig har inga skor på fötterna	2843	1036	1/17
218 *Hon hade en ny ring i fingret	3159	1282	17/2
221 Olle mötte Karin på tåget	2967	1431	19/1
224 Erik hade en ring på fingret	2425	860	14/7
Control sentences			
203 Vi måste vika av åt höger	3632	1297	11/8
207 Brevet är skrivet på finska	2815	1118	15/5
209 *Jag såg en polis att stå på gatan	2995	942	5/16
216 *Saken beror av din syster	2712	874	10/11
219 Flickan såg Emil gå ut	2888	1229	16/4
220 *Ni ska vika av till vänster	3560	1376	9/10
222 *Boken är skrivet på svenska	2467	783	13/8
223 De beror på hennes föräldrar	2206	716	20/1

TABLE 4 Mean RTs and judgments for Finnish L1 sentences by Finns (N=21).

Target Sentences	Mean	SD	Judgment Acceptable/ Unacceptable
302 *Tytöllä oli kaunis hattu päällä	1934	552	1/20
304 *Tiesin Pirjon asuvan Oululla	1941	398	0/21
306 *Hän otti sormuksen Pian sormelta	1928	552	1/20
307 Pojalla oli musta pipo päässä	1882	687	19/2
308 *Maija halusi riippua köydellä	2687	940	2/19
310 *Hän matkusti heti ulkomaihin	1797	1012	0/20
311 *Lehdet oksalla olivat kuolleet	3161	1257	4/16
312 Kuulin hänen viihtyvän Porissa	1981	729	21/0
314 Vedin sormuksen Liisan sormesta	1810	514	20/1
315 Lapset aikoivat riippua köydessä	2664	1018	14/7
317 *Katossa oli korkea savupiippu	3001	1573	6/12
318 *Tarja piti äitiä lujasti kädessä	2155	466	1/20
319 Lahtinen lähti eilen ulkomaille	1717	274	21/0
320 Lehdet oksassa olivat jo kuolleet	3061	1277	10/11
323 Katolla oli harmaa antenni	2223	586	21/0
324 Lapset puristivat Raisaa kädestä	2029	554	21/0
Control Sentences			
301 Hän on asunut Turussa vuodesta 1962	2119	517	21/0
303 *Pekka ei ollut syönyt omenan	1979	428	0/21
305 Ville myönsi halunsa ostaa auto	3142	1586	5/13
309 Kaisa ei ollut lukenut kirjaa	1935	485	21/0
313 *Myönsin haluni kirjoittaa kirjan	3100	1365	8/11
316 *Annoit käskyn ottaa pojan kiinni	2536	931	11/9
321 Annoin käskyn ottaa rosvo kiinni	2627	1412	16/3
322 *Mirja asuu Lapissa vuodesta 1980	2341	1124	1/20

Appendix 3

LIST OF TEST SENTENCES AND THE LEARNERS' RESPONSES

TABLE 1 Learners' mean RTs and judgments on English L2 sentences (N=66).

	Mean RT	SD	Judgment (Acceptable/ Unacceptable)
101 *George made her to steal it	2259	629	45/21
102 Jane made me drink it	1907	534	65/1
103 *He heard Jeff to come in	2266	803	22/44
104 She heard Mary go out	2043	616	52/14
105 *They watched him to go in	2107	655	36/30
106 She watched her playing outside	2306	657	47/19
107 *The judge ordered him sit down	2090	732	64/1
108 The major ordered us to stand up	2173	668	63/2
109 *Bill asked Jane dine with him	2220	576	12/54
110 Bob asked Betty to walk with him	2143	790	64/1
111 *He drink tea every evening	2005	539	7/59
112 She makes coffee every day	1882	728	64/1
113 *The boys loves their little sister	1962	550	8/58
114 The girls hate their little brother	2498	703	55/11
115 *My father have two brothers	1889	541	8/58
116 My mother has three sisters	1946	488	65/1
117 *He left it to Jills room	2407	900	29/36
118 She left it in Jacks car	2082	595	64/2
119 *The train arrived to the station	2190	585	26/39
120 The plane arrived at the airport	1892	541	58/8
121 *She thought her hair was in fire	2391	753	43/22
122 They said our house was on fire	2190	585	63/3
123 *There was no handle in the door	2327	576	44/22
124 There was no lock on the door	2223	764	56/9
125 *They wanted to sit in Jill's table	2448	653	28/38
126 She liked to sit at Joe's table	2290	555	56/10
127 *Jill had a ring in her finger	2303	708	45/21
128 Jim had a ring on his finger	2362	862	45/19
129 *He lived on the 12th century	2266	680	23/43
130 She died in the 8th century	2446	686	56/10
132 John is interested in cars	1796	618	64/2
133 *The price depends of it	1970	563	16/50
134 The answer depends on it	1827	426	64/2
135 *They did fun of him	2048	618	18/48
136 He made fun of me	1743	549	61/5
137 *I like of biscuits	1485	405	2/64
138 She likes chocolate	1491	943	62/2
139 *Jack Jones is born in 1952	2061	539	10/56
140 Bill Smith was born in 1940	1838	672	64/1

TABLE 2 Learners' mean RTs and judgments on Finnish L1 sentences (N=66). The location of the error has been indicated by an asterisk. The errors have not been translated into English.

	Mean RT	SD	Judgment (Acceptable/ Unacceptable)
201 *Pekka pisti hänet lukea Pekka made him/her *read	1500	377	0/66
202 Maija pisti minut tiskaamaan Maija made me do dishes	1876	595	62/4
203 *Me kuulimme hänen puhua We heard him/her *speak	1315	345	2/64
204 He kuulivat hänen laulavan They heard him/her singing	1430	374	65/1
205 *Mies näki sinun itkeä The man saw you *cry	1340	305	2/64
206 Tyttö näki hänen maalaavan The girl saw him/her painting	1699	705	63/2
207 *Ari käski häntä kiivetä ylös Ari told him/her *climb up	2219	934	22/42
208 Mika käski meitä tulemaan alas Mika told us to come down	1946	734	61/4
209 *Hän kävelivät töistä kotiin He/she walked (*Pl.) home from work	1626	439	43/23
210 Hän juoksi kotoa kouluun He/she ran from home to school	1726	542	65/1
211 *Te lauloi kauniin laulun You (Pl.) sang (*Sg.) a beautiful song	1558	504	6/60
212 Sinä ostit uuden paidan You bought a new shirt	1544	397	64/2
213 *Hän jätti sen Kaisan pöydällä He/she left it *on Kaisa's desk	1698	447	11/55
214 Hän jätti sen Mikon autoon He/she left it in Mikko's car	1610	391	65/1
215 *Kone saapui lentokentällä The plane arrived *at the airport	1527	355	6/59
216 Bussi saapui asemalle The bus arrived at the station	1471	713	64/1
217 *Näin kotini olevan tulella I saw my house was *on fire	1739	426	1/65
218 Luulin takkini olevan tulessa I thought my coat was on fire	1831	709	63/2
219 *Chopin eli viime vuosisadassa Chopin lived *in the last century	1838	633	2/63
220 Goya kuoli viime vuosisadalla Goya died in the last century	1876	683	64/1
221 *Äiti on ylpeä sinuun Mother was proud *of you	1373	404	2/64
222 Isä on ylpeä heistä Father is proud of them	1465	434	64/2

COMMUNICATIVE ANXIETY QUESTIOINNAIRE

Seuraavassa on joukko englannin kieltä ja yleensä kielenkäyttötilanteita koskevia väitteitä. Merkitse janalle rastilla (X), miten hyvin kukin väite pitää paikkansa Sinun kohdallasi. Merkitse rasti sitä lähemmäksi kohtaa samaa mieltä, mitä paremmin väite vastaa omaa mielipidettäsi tai kokemustasi. Mitä vähemmän väite vastaa omaa mielipidettäsi tai kokemustasi, sitä lähemmäksi kohtaa eri mieltä merkitse rasti. Lopussa on lisäksi neljä väitettä, jotka koskevat mielipidettäsi englannin taidostasi. Merkitse rasti sitä lähemmäksi janan päätä, mitä paremmin se vastaa arviotasi englannin taidostasi.

Esimerkki: Jos olet sitä mieltä, että allaoleva väite Luen englnnninkielisiä lehtiä usein pitää kohdallasi melko hyvin paikkansa (esimerkiksi tunnet lukevasi englanninkielisiä lehtiä melko usein), voit merkitä sen esimerkiksi seuraavasti:

0. Luen englanninkielisiä lehtiä usein.
samaa mieltä _____ eri mieltä

1. Koulussa minulla oli aina englannin tunnilla jotenkin epävarma olo.
samaa mieltä _____ eri mieltä

2. Äännän englantia huonosti.
samaa mieltä _____ eri mieltä

3. Minusta ei tunnu lainkaan ahdistavalta, jos en pysty ilmaisemaan itseäni englanniksi juuri sillä tavalla kuin alunperin halusin.
samaa mieltä _____ eri mieltä

4. Haluan karttaa kaikkia epäselviä tilanteita.
samaa mieltä _____ eri mieltä

5. En välitä, vaikka tekisin virheitä puhuessani englantia.
samaa mieltä _____ eri mieltä

6. En koskaan ajattele, miltä kuulostan toisten mielestä puhuessani englantia.
samaa mieltä _____ eri mieltä

7. Englannin tunneilla sain enemmän kielteistä kuin myönteistä palautetta.
samaa mieltä _____ eri mieltä

8. Pystyn ilmaisemaan englannin kielellä mitä haluan.
samaa mieltä _____ eri mieltä

9. Pelkään kuulostavani hassulta puhuessani englantia.
samaa mieltä _____ eri mieltä

10. Englannin tunneilla kiinnitettiin mielestäni liian paljon huomiota virheisiin.
samaa mieltä _____ eri mieltä

11. Puhun mielelläni, kun kaikki kuuntelevat minua.
samaa mieltä _____ eri mieltä

12. En jännitä koskaan puhuessani englantia.
samaa mieltä _____ eri mieltä

13. Minua pelottaa nousta puhumaan kaikkien muiden edessä.
samaa mieltä _____ eri mieltä

14. Pelkään tekeväni virheitä, kun puhun englantia.
samaa mieltä _____ eri mieltä

15. Sain mielestäni tarpeeksi kiitosta ja myönteistä palautetta englannin tunneilta.
samaa mieltä _____ eri mieltä

16. Minusta tuntuu ahdistavalta, kun en löydä oikeita sanoja puhuessani englantia.
samaa mieltä _____ eri mieltä

17. Nousen mielelläni pitämään puheen yhteen kokoontuneelle joukolle.
samaa mieltä _____ eri mieltä

18. Olen ujo ihminen.
samaa mieltä _____ eri mieltä

19. Ymmärrän englanninkielistä puhetta:
en juuri lainkaan _____ erittäin hyvin

20. Puhun englantia:
en juuri lainkaan _____ erittäin hyvin

21. Ymmärrän kirjoitettua englanninkielistä tekstiä:
en juuri lainkaan _____ erittäin hyvin

22. Osaan kirjoittaa englanniksi:
en juuri lainkaan _____ erittäin hyvin

YHTEENVETO

Kieliopillisuusarviot ja reaktioaikamittaus vieraan kielen tutkimusvälineenä

1 Tausta ja tavoitteet

Tutkimus käsittelee erityisen kielellisen tehtävätyypin, kieliopillisuusarvioiden ja reaktioaikamittauksen yhdistelmän käyttöä vieraan kielen tutkimuksessa. Kieliopillisuusarviolla tarkoitetaan kielentutkimuksessa kielenkäyttäjän, joko syntyperäisen kielenpuhujan tai kielenoppijan, tekemää arviota jonkin kielellisen ilmauksen kieliopillisuudesta. Tässä työssä käytetyt ilmaisut olivat kaikki erikielisiä lauseita, joista esimerkkinä seuraava englanninkielinen epäkieliopillinen lause **She made her to steal it.* Tutkimusasetelmaan kuuluu myös kielenpuhujien reaktioajan mittaus. Lauseiden esitys on ollut tarkasti kontrolloitu joko erityisen diaheittimeen kytketyn mittarin avulla tai sittemmin tietokoneen avulla. Laitteiden ja ohjelmistojen avulla on myös kielenpuhujien vastaus ja siihen kulunut aika pystytty tallentamaan.

Kieliopillisuusarvioita on käytetty etenkin puhtaasti lingvistisessä, universaalikielioppia viitekehyksenä käyttävässä vieraan kielen oppimisen tutkimuksessa. Tällaisen tutkimuksen tarkoituksena on saada tietoa kielenoppijan intuitiivisesta tiedosta omasta kielijärjestelmästään eli kielellisestä kompetenssista. Kielenoppimisessa on avainasemassa universaalikielioppi, joka toimii kielellisen tiedon omaksumisen ohjaajana. Kieliopillisuusarviot on tämän tutkimusperinteen piirissä pitkään nähty tiedonkeruumenetelmänä, jolla on erikoisasema intuitiivisen kielitiedon kuvaajana. Tämä perustuu Chomskyn (1965) esittämään näkemykseen kielenpuhujien kompetenssiin kuuluvasta kyvystä erottaa kieliopilliset lauseet epäkieliopillisista. Reaktioaikamittaus on puolestaan psykometrisen tutkimuksen apuväline. Tälle tutkimusperinteelle on tyypillistä nähdä kielenoppija tiedonprosessoijana. Kielenoppimista pidetään usein samanlaisena prosessina kuin minkä tahansa kognitiviisen taidon oppimista. Jyväskylän kontrastiivisessa projektissa kieliopillisuus-arvioita ja reaktioaikamittausta on hyödynnetty kielenoppijan kielellisen suorituksen eli performanssin kuvauksessa. Tällaiset arviot eivät tarkasti ottaen ole kieliopillisuusarvioita vaan kieliopillista hyväksyttävyyttä koskevia arvioita; hyväksyttävyys käsitteenä liittyy ennen kaikkea kielenkäyttötilanteeseen ja se on aina suhteellista.

Tutkimuksen lähtökohtana on jännite, joka vallitsee kielentutkimuksessa kompetenssin ja performanssin tutkimuksen välillä. Etenkin kieliopillisuusarvioiden erityisasema on asetettu kyseenalaiseksi (ks. esim. Birdsong 1989a, Ellis 1991). Viime vuosina myöskin universaali-kielioppiin nojautuvat kielentutkijat ovat ryhtyneet keskustelemaan kieliopillisuusarvioiden suhteellisuudesta. Viime vuosina niitä on ryhdytty pitämään vain yhtenä esimerkkinä kielellisistä päätöksenteko- tai ongelmanratkaisuprosesseista. Jyväskylän kontrastiivisessa projektissa kieliopillisuusarviot on aina nähty ennen muuta kielellisen performanssin heijastajina (Lehtonen ja Sajavaara 1985). Reaktioaikamittauksen avulla on pyritty luomaan valoa kielellisten prosessien automaatistumiseen; automaatistumista pidetään sujuvan kielellisen suorituksen tunnus-

merkkinä. Sujuva kielellinen suoritus on taas vieraan kielen oppimisprosessin päätavoitteita. Tutkimuksessa keskitytäänkin kuvaamaan kognitiivisia kielenoppimisteorioita ja kielenkäyttömalleja.

Andersonin (1979, 1980, 1983) kehittämä ACT* (* lausutaan 'star') teoria pyrkii kuvaamaan taitojen oppimista kognitiiviselta kannalta. Vieraan kielen oppimiseen sitä ovat soveltaneet mm. Dechert ja Raupach 1987, Faerch ja Kasper (1986, 1987) sekä Hulstijn (1989, 1990). Keskeisessä asemassa Andersonin teoriassa on deklaratiivinen ('mitä') tieto ja prosedu-raalinen ('kuinka') tieto. Taitoa, siis myös kielitaitoa opittaessa deklaratiivinen tieto muuntuu erilaisten vaiheiden kautta proseduraaliseksi tiedoksi, jolloin tiedon saatavuus ja käyttö nopeutuu ja tulee vähitellen automaattiseksi. Automaatistumisesta on sinällään lukuisia teoriota. Usein tehdään ero automaattisten ja kontrolloitujen (joita kutsutaan myös ei-automaattisiksi) prosessien välillä (Shiffrin ja Schneider 1977). Tämä ero on laadullinen: automaattisia prosesseja pidetään nopeina, vaivattomina, vähän huomiota vaativina ja kerran omaksuttuina vaikeasti muutettavina, kun taas kontrolloituja prosesseja pidetään hitaampina ja prosessointikapasiteettiä vaativina. Prosessin automaatistuessa sen käsittämä säännöstö, algoritmi, nopetuu. Toisenlaista näkemystä edustaa ns. tapausteoria (Logan 1988). Tämän teorian mukaan automaattisilla ja kontrolloiduilla prosesseilla on edelleenkin laadullinen ero, mutta automaattisten prosessien nopeuden ajatellaan perustuvan siihen, että ne haetaan suoraan muistista.

Tutkimuksessa on kuvattu kielenoppijoiden kielellistä suoritusta sellaisissa lauseissa, joiden arviointiin liittyy mahdollisuus käyttää joko äidinkieleen tai englantiin perustuvaa kielitietoa arvion perustana. Kielten väliseen vaikutukseen eli kielelliseen siirtovaikutukseen (transfer) liittyvät tekijät ovat siksi erityisasemassa. Tällaisia tekijöitä ovat mm. kielten välinen välimatka: kielenoppijalle on oppimisprosessin aikana tai jo sitä ennen muotoutunut käsitys siitä, miten lähellä tai kaukana hänen oma äidinkielensä ja kohdekieli ovat rakenteellisesti (Kellerman 1977). Esimerkiksi suomalainen kielenoppija mitä todennäköisimmin pitää suomea rakenteellisesti hyvin erilaisena kuin englantia. Tämä puolestaan vaikuttaa kielenoppijan haluk-kuuteen nojautua äidinkieleensä englantia opittaessa ja/tai prosessoitaessa: kielenoppija ei ole yhtä valmis siirtämään äidinkieleen perustuvia rakenteita vieraaseen kieleen.

Kieliopillisuusarvioihin liittyviä metodologisia näkökohtia on viime vuosina pohdittu melko runsaasti (Birdsong 1989a, Ellis 1991, Hedgcock 1993, Cowan ja Hatasa 1994, Gass 1994). Kieliopillisuusarvioihin vaikuttavat tekijät voidaan ryhmitellä kolmeksi keskeiseksi muut-tujaryhmäksi. Näistä ensimmäinen koskee kielellisen ärsykkeen laatua: etenkin se millaisen virheen epäkieliopillinen lause sisältää vaikuttaa kielenoppijan arvioon. Toisaalta arviointiprosessiin vaikuttavat arvioijan kognitiiviset ja affektiiviset ominaisuudet, esimerkiksi se, tunteeko hän ahdistuneisuutta tekemistään kielellisistä virheistä. Kolmantena tekijänä arviointiprosessiin vaikuttaa yleensäkin kokeelliseen tutkimukseen liittyvät tilanneym. tekijä: millaiset ohjeet on annettu, miten ne on annettu, arvioijan vireystila yms. Tällaiset tekijät on suhteellisen helppo sijoittaa esim. Lehtosen ja Sajavaaran (1980) malliin kielellisen viestin prosessoinnista. Lehtonen ja Sajavaara näkevät erilaisten kognitiivisten ja affektiivisten tekijöiden vaikuttavan kielellisen viestin prosessointiin. Samankaltaisia näkemyksiä on viime aikoina esitetty kielitaidon mittauksen tutkimuksen piirissä: Bachman ja Palmer (1996) esittävät joukon kognitiivisia ja affektiivisia tekijöitä, jotka vaikuttavat kielenkäyttöön testitilanteessa.

2 Kokeellinen osa

Tutkimus koostuu kolmen kokeen sarjasta. Näistä ensimmäisen tuloksia on lähemmin tarkastelu ja aiemmin (Alanen 1987, 1991). Kussakin kokeessa koehenkilöitä pyydettiin arvioimaan englanninkielisten ja suomenkielisten lauseiden kieliopillista hyväksyttävyyttä. Koe 1 sisälsi myös saksan- ja ruotsinkielisiä, Koe 2 ruotsinkielisiä lauseita.

2.1 Koe 1

Tässä kokeessa tarkoituksena oli tutkia lähemmin suomalaisten englanninoppijoiden taipumusta siirtää englannin rakenteita heikommin hallittuun vieraaseen kieleen, mitä tarkoitusta varten saksan- ja ruotsinkielisiin lauseisiin oli rakennettu englantiin pohjautuvia virheitä. Koe koostui kaiken kaikkiaan 91 englanninkielisestä, 34 saksankielisestä ja 29 ruotsinkielisestä lauseesta. Näiden lisäksi siirtovaikutusta pyrittiin tutkimaan erityisen käännöksen oikeellisuutta koskevan arviontitehtävän avulla: koehenkilöille näytettiin lausepareja, joista ylempi oli aina englanninkielinen lause ja alempi lause sen saksan- tai ruotsinkielinen käännös. Käännöksiin oli upotettu englantiin pohjautuvia virheitä. Koe sisälsi myös suomenkielisten lauseiden arviointitehtävän ja englanti-suomi-käännös-arviotehtävän. Koehenkilöinä oli suomalaisten englanninopiskelijoiden (N=20) lisäksi 15 syntyperäistä englanninpuhujaa, 13 syntyperäistä saksanpuhujaa, jotka koostuivat maassa oleskelevista saksan kielen opettajista, ja 20 saksalaista englanninopiskelijaa.

Tuloksista kävi ilmi, että syntyperäisen kielenpuhujat olivat merkittävästi nopeampia arvioidessaan äidinkielisiään lauseita kuin kielenoppijat; he myös osuivat merkittävästi useammin oikeaan. Lisäksi tulokset antoivat viitteitä siitä, että suomalaiset kielenoppijat nojautuivat vahvempaan englannin kieleen saksankielisiä lauseita arvioidessaan: mm. englantiin pohjautuvat passiivirakenteet hyväksyttiin saksassa. Suomalaisten ja saksalaisten englanninoppijoiden reaktiot epäkieliopillisiin englanninkielisiin lauseisiin erosivat myös toisistaan riippuen siitä, sisälsivätkö ne mahdollisesti jompaan kumpaan kieleen pohjautuvan rakenteen. Koe paljasti myös mielenkiintoisia systemaattisia yksilöllisiä eroja kielen-oppijoiden välillä: jos koehenkilö oli hidas arvioidessaan suomenkielisiä lauseita, hän oli hidas myös englannin- tai muunkielisiä lauseita koskevissa arvioissa. Kielitaidon taso näytti myös vaikuttavan koehenkilöiden suoritukseen: syntyperäisten kielenpuhujien paremmuus on jo mainittu, mutta tämän lisäksi paljastui, että saksaa yliopistossa opiskelevat (N=10) suomalaiset olivat merkittävästi nopeampia saksankielisissä tehtävissä kuin saksaa opiskelemattomat (N=10). Nämä tulokset saattavat tuntua itsestäänselviltä; on kuitenkin huomattava, että kieliopillisuusarvioiden luotettavuudesta on esitetty runsaasti epäilyksiä. Tulokset näyttivät kuitenkin ainakin näiden lauseiden osalta varsin johdonmukaisesti heijastavan paremman kielitaidon tuomaa sujuvuutta.

2.2 Koe 2

Edellisessä kokeessa oli havaittu johdonmukaisia eroja yksittäisten kielenoppijoiden suorituksessa. Tässä kokeessa haluttiin lähemmin tarkastella kieliopillisuusarvioiden nopeuden ja oikeellisuuden suhdetta joukkoon kognitiivisia ja affektiivisia tekijöitä. Näitä tekijöitä olivat kenttäriippuvuus/ riippumattomuus, vieraskielisiin viestintätilanteisiin liittyvä ahdistuneisuus, verbaalinen tai visuaalinen ajattelutyyli ja äidinkieltä koskeva päättelykyky. Kenttäriippuvuuden mittarina käytettiin Witkinin (Witkin et al. 1971) työtovereidensa kanssa luomaa testiä (Embedded Figures Test), jossa mitataan se aika, joka koehenkilöltä kuluu yksinkertaisen kuvion erottamiseen mahdollisesti hämäännyttävästä taustasta. Mitä nopeampi koehenkilö tehtävässä on, sitä riippu-mattomampi kentästä hän on. Tällaisia

henkilöitä pidetään yleensä analyyttisimpinä kuin tehtä-vässä hitaampia kenttäriippuvaisia yksilöitä. Kommunikatiivisen ahdistuneisuuden mittaamiseen käytetiin kyselykaavaketta (Lehtonen, Sajavaara ja Manninen 1985) kuten myös verbaalisen ja visuaalisen ajattelutyylin kuvaamiseen (Paivio ja Harshman 1983, Spitzer 1986). Äidinkieltä koskevan päättelykyvyn mittarina käytettiin Työvoimaministeriön ammatinvalinnanohjauksessa käyttämää suomenkielistä aukkotekstiä. Kieliopillisuusarvioita pyydettiin englannin-, ruotsin- ja suomenkielisistä lauseista (k=24), jotka sisälsivät tietyntyyppisisiä prepositiorakenteita (k=16). Kohdelauseiden lisäksi kokeeseen kuului joukkoverrokkilauseita (k =8). Koehenkilöinä oli 21 suomalaista englanninopiskelijaa ja viisi syntyperäistä englanninpuhujaa.

Tulo-momenttikorrelaatioiden laskeminen ei paljastanut merkittäviä korrelaatioita reaktioaikatehtävän ja kognitiivisten ja affektiivisten muuttujien välillä muiden kuin vieraskieliseen viestintätilanteeseen liittyvän ahdistuneisuudeen osalta. Kentäriippuvuus/riippumattomuus, ajattelutyylin verbaalisuus ja äidinkieltä koskeva päättelykyky eivät olleet missään suhteessa arviointien nopeuden tai tarkkuuden kanssa. Syntyperäiset englanninpuhujat olivat jälleen parempia kieliopillisuusarvioissaan; tosin ero nopeudessa ei ollut tällä kertaa niin suuri kuin ensimmäisessä kokeessa. Aivan kuten ensimmäisessä kokeessa, kielenoppijoiden suoritusajat eri kielissä korreloivat voimakkaasti keskenään.

2.3 Koe 3

Kolmannen kokeen tarkoituksena oli saada selville kielenoppijoiden kielitaidon tason ja reaktioaikatehtävän suhde. Koehenkilöinä oli 66 suomalaista englanninopiskelijaa. Koehenkilöitä pyydettiin arvioimaan englanninkielisten (k=39) ja suomenkielisten (k=22) lauseiden kieliopillista hyväksyttävyyttä ensin reaktioaikatehtävien avulla ja sitten kirjallisessa muodossa. Suomenkielisten ja englanninkielisten lauseiden sisältämät virheet oli lisäksi pyritty luomaan mahdollisimman samankaltaisiksi. Kielitaidon mittarina käytetiin mukaeltua pääsykoetta. Koehenkilöiden vieraskieliseen viestintätilanteeseen liittämää ahdistuneisuutta mitattiin kyselykaavakkeella (Lehtonen, Sajavaara ja Manninen 1985). Koehenkilöiden yksinkertainen reaktionopeus mitattiin myös.

Tulokset paljastivat heikon mutta tilastollisesti merkittävän käänteisen korrelaation reaktioajan ja kielitaitotestin pistemäärän välillä. Mitä nopeampi kielenoppija oli arvioissaan, sitä korkeampi oli hänen pistemääränsä. Jonkin verran selvempi suhde paljastui kielitaitotestin kieli-oppirakenteiden hallintaa koskevan osion ja vastausnopeuden ja -tarkkuuden välillä. Mitä korkeampi pistemäärä kielioppiosiossa koehenkilöllä oli, sitä tarkempi hän oli arvioissaan. Tulokset viittaisivat siis siihen, että reaktioaikatehtävän ja kielitaidon välillä olisi yhteys. Koska korrelaatiokertoimet olivat melko alhaiset, niiden tulkinnassa on tilastollisesta merkitsevyydestä huolimatta syytä olla varovainen. Tulokset ovat siksi enemmänkin suuntaa-antavia. Selvempi suhde ilmeni kielitaidon ja kielenoppijoiden tietyntyyppisiä epäkieliopillisia lauseita koskevien arvioiden välille. On kuitenkin merkillepantavaa, että kokeessa ei ollut havaittavissa nopeuden mukaantuomaa epätarkkuutta. Koehenkilöjoukon 16 nopeinta ei ollut arvioissaan merkittävästi huonompi kuin 16 hitainta. Tämänkaltainen nopeuden ja tarkkuden yhdistelmä on tyypillistä sujuvalle ja taitavalle suoritukselle.

Kielenoppijoiden yksilöllisessä suorituksessa paljastui myös samantyyppisiä eroja kuin aikasemmissa kahdessa kokeessa. Koehenkilöt, jotka olivat hitaampia arvioimaan suomenkielisiä lauseita, olivat hitaampia myös englanninkielisten lauseiden arvioinnissa.

Yksilöllisten erojen selville saamiseksi tuloksiin sovellettiin ryhmittelyanalyysia. Ryhmittelyn tuloksena koehenkilöistä muodostui neljä erillistä ryhmää: 1. koehenkilöt (N=18), joilla oli muita korkeampi kielitaitotestin pistemäärä; 2. koehenkilöt (N=8), joiden kielitaito ja tarkkuus reaktioaikatestissä oli huonompi kuin muiden; 3. koehenkilöt (N=12), jotka olivat arvioissaan hitaampia kuin muut; ja 4. koehenkilöt (N=21), jotka olivat arvioissaan tarkempia kuin muut Kartoitettaessa syntyneiden ryhmien taustamuuttujia paljastui, että Ryhmä 1 koostui suureksi osaksi sellaisista opiskelijoista, jotka olivat oleskelleet keskimääräistä pitempään englantia puhuvassa maassa. Ryhmä 4 puolestaan sisälsi selvästi enemmän kielenoppijoita, jotka olivat viettäneet vain vähän aikaa ns. luonnollisessa kielenoppimisympäristössä. Silti juuri Ryhmä 4 oli muita tarkempi ja nopeampi reaktioaikatehtävissä. Tämä tulos yhdessä sen seikan kanssa, että keskimääräisten äidinkielisten ja vieraskielisten arviointiaikojen välillä oli voimakas korrelaatiosuhde, antaisi aiheen olettaa, että reaktioaikatestiin olisi kielitaidon lisäksi liittyneenä myös metalingvistisiä taitoelementtejä.

3. Tulosten merkitys

Kokeiden tuloksissa paljastui, että kielitaidon ja reaktioaikatehtävän välillä on ilmeinen yhteys. Erityisen selvä tämä yhteys on tarkasteltaessa syntyperäisten kielenpuhujien ja kielenoppijoiden suorituksia: syntyperäiset kielenpuhujat olivat yleensä ottaen merkittävästi nopeampi ja tarkempia arvioissaan. Yhteydet kielitaitotestin ja reaktioaikatehtävän välillä eivät olleet yhtä selkeät. Reaktioaikatehtävä näyttää ennen kaikkea paljastavan mielenkiintoisia eroja pintapuolisesti samantasoisten kielenoppijoiden kielellisessä suorituksessa.

Reaktioaikatehtävät sellaisena kuin niitä sovellettiin tässä tutkimuksessa ovat tyypillisiä kokeelliselle, usein laboratoriossa suoritettavalle tutkimukselle. Reaktioaikatehtävät jakavat myös tämäntyyppisen tutkimuksen heikkoudet: kielen käyttöympäristö ei ole luonnollinen; kielelliset ärsykkeet koostuvat yksittäisistä lausetason ilmiöistä, jotka esitetään ilman kielellistä ja sosiaalista kontekstia. Tulosten soveltamiseen koetilanteen ulkopuolelle on niin muodoin suhtauduttava varoen. Toisaalta voidaan tämäntyyppisellä tutkimuksella usein vangita tietoa, joka luonnon-mukaisemmassa kielenoppimis- ja käyttöympäristössä saattaa joskus jäädä huomaamatta. Kummankin tutkimussuunnan voidaan parhaassa tapauksessa katsoa täydentävän toisiaan.

Reaktioaikatehtävän suhteesta kielellisiin tuotoksiin olisi hyvä saada lisää tietoa.; nythän sekä tehtävä itse että kielitaidon mittari olivat lähinnä reseptiivisiä luonteeltaan. Reaktioaikatehtävän suhteuttaminen suulliseen tuottamistehtävään saattaisikin tarjota mielenkiintoista lisätietoa kielellisen tiedon luonteesta ja käytöstä.