Plural Processing in Native Speakers and Learners of English: Challenging the Notion of Strictly Grammatical Plural Processing

Dissertation zur Erlangung des Doktorgrades in
der Philologie an der Universität Paderborn
Gutachter: Prof. Dr. Manfred Pienemann
Prof. Dr. Gisela Hakansson (Lunds universitet)

von
MARIE-LOUISE POSCHEN
geboren am 8. Mai 1973 in Marburg / Lahn

Plural Processing in Native Speakers and Learners of English: Challenging the Notion of Strictly Grammatical Plural Processing

“Leven is meervoud van lef”

Loesje

Acknowledgements

This dissertation has been part of my life for many years. A number of people have lived through it with me and have supported me and my work in various stages. For this I would like to thank them.

First and foremost I would like to thank Professor Dr. Manfred Pienemann for taking me on as a doctoral student and for his willingness to supervise a dissertation on a topic that proved to be less straightforward than we both might have wished.

I owe a debt of gratitude to the large number of pupils, teachers, and British military personnel who provided me with such reliable data. I would also like to thank the teachers and administrators at the King´s School Gütersloh, the Reismann Gymnasium Paderborn, the Gymnasium Delbrück, and the Electronic Unit at the Barker Barracks in Paderborn. I will never forget the sight of some 20 soldiers in uniform marching to the computer lab at the University of Paderborn in order to take part in the experiment as part of their exercise drill. The cooperation with all participants and their superiors proved to be the most enjoyable aspect of the entire PhD project.

The most influential person involved in the process of my dissertation, and the person who brought colour into my life in Paderborn, was Tanja Ruthenberg, who has been my friend and colleague from the first moment we met. Tanja and I shared many frustrating, important, dull, joyful, and hilarious moments. Indeed, she is the only person who knows what I am like under all of those circumstances. Sharing an office, Tanja and I laughed together, cried together, and supported each other through the crises in our daily lives; and we both grew in the process over the years. Had it not been for Tanja, I might have abandoned this endeavour shortly after having undertaken it. I certainly would have learned less without her, and I definitely would not have had such a good time. And without Tanja, I would never have met my husband! Tanja, I would like to thank you for your friendship and for being the person you are.

Thanks a million to Norbert Hagemann who was never too busy or too tired to solve my statistical or formatting problems. Whenever I asked him, he immediately sat down and explained everything patiently. Norbert, you are a great statistician, formatter, and above all a great person!

Special thanks must go to my teacher, mentor, and friend Wander Lowie, who was an enormous source of inspiration in the many supervising sessions in Groningen. Wander, you were such a big help with the computer programme E-Prime at a time when I was utterly lost. Your many useful tips helped me strengthen the theoretical part of my dissertation. Thank you too for proofreading those chapters and for taking so much time to help your former student. My sessions with you in Groningen inspired me and filled me with the confidence and conviction I needed to see this project through. They also reminded me of how much fun it could be to work and exchange ideas with others in an academic environment. In this vein I would also like to mention Floor Buschenhenke and Jacqueline van der Poel who contributed so much to those sessions. Jacqueline and Wander, you were excellent hosts: thank you for sharing your culinary delights with me!

I would very much like to thank Oliver Scholle, who was an enormous help programming a program to match the frequencies of the stimulus words from the CELEX word list. Without Olli, I would never have been able to test what I set out to test. Thank you, Olli, for selflessly spending so many weeks helping me. I know you also had a lot of things to do. I really appreciate what you did.

I would like to thank Vijaya Kohli for being a friend, source of inspiration, source of encouragement, and simply for believing in me. Apart from sharing a love of balcony plants, puppadams, and dawn walks around the Padersee, Viji always showed interest in me and my project, and believed it would be completed. Thank you, Viji, for so many wonderful hours and for giving me the opportunity to return to Paderborn to teach and to use your office.

My thanks also to Professor Dr. Günter Rohdenburg for being such a kind spirit with a great sense of humour. Thank you for providing plenty of opportunities to chat and for inviting me to teach in your class.

I would also like to thank Anke Lenzing, Jana Roos, and Gabriela Hoffmann for the help you provided. Even though we were not always in the same city, it was heart-warming that we were able to help each other with our dissertations and to create a mutually fruitful study environment. Anke and Gabi, thank you also very much for proofreading the conclusion to my dissertation.

I should not forget to thank Helen Snively. You were a really good 'thesis therapist,' Helen! Thank you so much for all your editing suggestions, your help, your patience, and your emails. Knowing that someone was reading my emails and following my thoughts actually made the process of completing the conclusion fun. The story of your role in editing helping me put the finishing touches on my thesis has become part of our family history. No future Christmas will pass without raising a glass to you.

This dissertation became such an all-consuming part of my life over the years that it was perhaps inevitable that I would share my trials and tribulations with my closest friends. Aly Jellema, Martin de Bruijn, and Ernest Ram are just a few of the friends who had to endure my constant complaining. I even cut a wonderful holiday short because I thought I had to work on my dissertation. What a mistake! I am also indebted to all the friends whom I have not mentioned, those who contributed to the success of my project by word and deed, and those who contributed simply by being there for me. At some point, every one of you heard the excuse that I could not keep in touch because I was so extremely busy with my dissertation. Please accept my apologies.

My great-grandma Ruth Köller and my grandma Inge Poschen always showed great interest in my work and supported me in their thoughts. Thank you so much! Grandma Ruth promised to be around to celebrate the completion of my thesis with me. She kept her promise before passing away at the age of 102.

I would like to thank my mum and dad for being there and for always supporting me. It was good to know that you always believed in the importance of my dissertation even though you had no idea what it was about. It was also good to know that it was completely irrelevant to you whether I wrote a dissertation or not. You might have been better off had I not researched and written this thing. It was not always very easy, and you were the first to hear. Thank you for the love you keep on giving.

Finally, I would like to thank my husband, Thomas. You played the decisive role in focusing my attention on completing the dissertation that I had neglected for many years. I cannot thank you enough for helping me close one chapter of my life and open a much more exciting new chapter. And thank you for the role you play in my life. In the end, Alva, Daan, and you made me see the dissertation in its proper perspective: its importance has faded completely compared to what you have given me.

1 *Introduction*

A remarkable skill in living organisms is the capacity to adapt the living body to the environmental reality. The very definition of a living organism is its constant exchange with its extraorganic reality. When an organism alters its behavior due to this constant exchange with the environment, it has learned to internalize knowledge about the world. Learning can be either conscious or subconscious. Not only may human beings learn subconsciously, they may also learn without any change in overt behavior. Learning may not always manifest itself in overt behavior. The acquisition of factual information or the acquisition of operational routines may, thus, not always be clearly visible.

> Learning is a process that can occur in the absence of overt behavior but its occurence can only be inferred by seeing changes in behavior. The overt behavioral changes, and the other changes that cannot be detected simply by the organism's overt behavior, all reflect alterations in the brain produced by learning. For this reason, although learning has been very fruitfully studied by purely behavioral techniques, many of the fundamental questions about learning will require direct examination of the brain (Kupfermann, 1985: p. 806).

This applies to all kinds of learning and to language learning in particular. Concerning language learning, it is important to distinguish between the acquisition of the first language and the learning of any subsequent language. This study will concentrate on second language learning only, i.e. any language learnt after the first language has been acquired. More specifically, this project studies second langauage acquisition in a limited context at a German school and compares the findings to native speakers´ language processing. Some of the questions related to language learning are very general and apply to both kinds of acquisition, some are specific to only one of the two. The general questions that guide this whole research project are the following: How is it that we actually learn? How does input actually lead to linguistic processing and to the subconscious acquisition of grammatical knowledge?

In our German school system we assume that language is learnable within a few hours every week over a period of a couple of years. Usually, school children learning English as a second language at grammar school have at least five years of compulsory language classes. It is expected that exposure to the teacher´s language input (usually provided by only one or two teachers) as well as explicit grammar teaching with rules and exercises as well as proficiency exercises will eventually result in advanced foreign language skills. It becomes difficult to assess whether pupils have actually acquired a foreign language in a near native way at the end of their school career if it is not the visible performance that may tell us something about the learning process. Students may perform brilliantly in proficiency exercises or they may have internalised several ways to produce fluent texts and may be flawless at writing or speaking, but what goes on in their heads remains unknown to the teacher. The actual mental processing of language may not be native-like at all. Moreover, second language learners performing flawlessly in language tasks may take significantly longer processing time than natives, they may also use different mental procedural skills than natives in order to operate in the second language. So processing time and processing procedure may be different from natives language users. Therefore, it is important to determine whether a change in language processing has taken place since learning can take place in the absence of overt behavior. At the same time, however, it can be assumed that in spite of changes in overt linguistic behaviour, changes in linguistic processing may not have taken place. For teachers at a German grammar school, the actual results in langauge tasks may suffice to assess a student´s language behaviour.

For the linguist who wants to study the nature of learning, the hidden processing procedures are vital. It is exactly this interest in the nature of processing and automatisation of simple structures in learner language, that this thesis set out to explore. The data provided here will yield new insights in the exact nature of the factors that influence a particular type of language processing and how they develop in language learners.

1.1 The present study

Learning a foreign language consists of implicit and explicit learning. Implicit learning occurs when a shift in linguitic behaviour takes place because a learner internalises a new linguistic item, for instance the past tense rule. Implicit learning may take place without any explicit awareness of such a rule. It may be the result of an automatisation process because of frequent usage of the regular past tense in the learner´s linguistic input. Explicit learning takes place if a learner uses the past tense rule because he has studied the rule itself and knows how to apply it to the verbs in question. In both cases, possible overgeneralisation of the past tense rule ("wented" instead of "went") shows that the learner has acquired the structural property in question. The question of implicit and explicit will not be further addressed in this study since the task in the experiment to follow cannot make any assumptions whatsoever about the implicit or explicit nature involved.

Apart from the distinction between implicit and explicit learning, there is another one that needs to be made: the distinction between lexical and grammatical learning. Learning a new language means learning new vocabulary. In this process, new words need to be memorised in a mental "dictionary" -the mental lexicon-, from which we may retrieve the appropriate words when necessary. Apart from the words, a language also has structural properties that often pose a problem for second language learners since these differ widely from language to language. When a new grammatical structure is introduced in the language learning classroom, it is often practised explicitly by means of a rule. The introduction of such a rule in teaching, say the noun plural rule, often follows the first plural usage in the language material in the textbook. Whether the teacher introduces the rule deductively or inductively is not of any interest here. But it is important to distinguish between lexicon and grammar. Grammar is related to structural properties that operate on the lexical items that we store in our mental lexicon. We apply the grammatical rules to the vocabulary that we know. For instance, English pluralisation of nouns can be explained by the simple rule of adding an "s" to the singular form of the word. For the purpose of classroom teaching, the explicit knowledge about a *plural rule* may be very helpful. An important question of this research project is how the learner processes such a rule when applying it. How does the learner arrive at a near-native command of such morphological or grammatical rules? This study will concentrate on the processing nature of plural nouns in language learners and native speakers of English. As stated above, learners may show correct overt linguistic behaviour when applying the plural rule but may not have acquired native-like linguistic processing skills. In order to find out whether they show native-like behaviour, learner groups will be compared with a native speaker group. However, the complete nature of plural processing is not completely understood for native speakers. This thesis sets out to explain which exact factors influence the processing of a seemingly simple plural rule. This understanding, in turn, should help to further develop theoretical models of language representation and language processing, which will be presented in the following chapters.

Any study of learners´ language processing attempts to find out the nature of the mental sub-processes involved and how it changes in the process of second language acquisition. When judging a language learner´s linguistic processing competence of the plural rule, it does not suffice to hear a learner apply the plural "s" correctly when speaking. Instead, the nature of the cognitive processes involved are of interest. These can be studied by examining the processing time, for instance. Such information may give some insight into how a learner generates this pluralised form from the knowledge of a simple singular word and a plural rule to be applied to that word.

1.1.1 Research methodology

The psycholinguistic research techniques used in the last couple of decades allow us to sketch a very accurate picture of the processes involved in speech production and comprehension. This so called on-line experimental work provides new insights into the actual processes at work while a person exercises certain linguistic tasks. Reaction time experiments, for instance, can resolve the cognitive process of language processing during the time course of the actual act of speech comprehension or production. This allows more accurate statements about the time order of such processes and therefore helps to understand the modelling of language cognition.

In this thesis, the processing of plural words will be analysed in terms of reaction time. Moreover, a model of plural processing will be suggested that incorporates a multitude of factors influencing the processing nature of English pluralisation. English plural inflection is seen as a multi-facetted processing event in contrast to the assumption that it is purely grammatical rule application, as it is understood in the language classroom. In this context, another question arises: are some properties of language such as the English plural of a clearly discernible grammatical nature, in the sense that learners always apply these *plural rules* to the word stems stored in their mental lexicon when processing pluralised words? Such a process implies that the rule is applied at a so-called *extra-lexical* stage and not while accessing the word stem from the mental lexicon (Perlmutter, 1988).

On the other hand, the plural may not be so grammatical at all and may also display some lexical[1] properties. Lexical, in this case, means that some of the fully pluralised forms might already be stored as complete chunks in the speaker´s mental lexicon. In that case it would be redundant to apply the plural rule. Various factors may influence the processing nature of pluralisation in the foreign and also in the native language. This thesis will be concerned with the processing mechanisms involved in the usage of the plural by native speakers and second[2] language learners, drawing from existing work on native language processing of inflections.

The experimental study presented in this thesis is developed to test second language learners´ reaction times in on-line processing of plural inflections by means of two reaction-time experiments. The hypothesis is that, opposed to the strict distinction between lexicon and grammar proposed and experimentally supported by various scholars (Perlmutter, 1988; Badecker & Caramazza, 1989; Schriefers, Friederici, & Graetz, 1992), a less strict separation between grammar and lexicon must be assumed at least for morphology. A gradual difference in inflectional processing -from strictly grammatical processing to more or less fully lexically inherent forms- should be mirrored in the phenomenon of learners´

1. The reference to grammatical and lexical properties refers to the Levelt (1989) model to be introduced in chapter 2

2. In this context, the second language is any language other than the native language and therefore covers the whole range of developmental stages. Accordingly, a bilingual is referred to as someone who speaks more than one language, regardless of the nature of acquisition.

reaction times gradually changing as their language develops. The determining factors for plural processing found in native speakers -such as frequency and dominance- are expected to appear only gradually in the language learners because they are highly related to the amount of language input the learners are exposed to.

In order to make claims about the nature of processing in language learners, a cross-sectional study of several learner groups of English as a second language has been carried out using two different experiments, to be described in Chapter Six. All learner groups are at different levels of language development and therefore they represent a cross section of learning, at least for the amount of input they have been exposed to in the classroom. For practical reasons I need to assume that the exposure is constant throughout the school years. Studying subjects at different developmental stages should shed light on the processes involved in lexical access and should also predict in what order these processes are acquired as well as which processing mechanisms are likely to be transferred. The learners in this study had different amounts of exposure to the target language and very limited input. The questions of this project particularly address the role of input since the learners and the native speakers take part in an experiment where word frequencies play an important role. Word frequencies are established by counting the number of appearances of a given word in a representative corpus of real language samples. As such, frequencies are based on native language usage and should not play a role for the beginning learner. The question that this study attempts to answer is the emergence point of such factors as frequency in learners´ processing development considering that they were taught in a classroom far away from native speaker input and frequencies.

As mentioned above, this study addresses some of the current issues about the nature and the systematicity of morphological processing procedures in natives and second language learners. It was already proposed that inflectional rules such as the plural rule are of a rather grammatical nature. This is suggested by several lexical access models to be introduced in the following chapters (e.g. Taft & Forster, 1975; Baayen, Dijkstra & Schreuder, 1997). These researchers argue that morphological processing of inflections is not intrinsic to the lexicon. Inflectional[3] morphemes are assumed to be processed extra-

3. The discussion about the nature of inflectional versus derivational morphemes will follow in a later section (5.1)

lexically and are consequently generated through grammatical rule application. The following issues are therefore central to this study:

It is the primary aim of this study to find evidence in native speakers that inflectional processing of the English plural does not follow the strict grammatical route. In Chapter 5, I will present arguments that make the strict separation between grammar and lexicon problematic to sustain. This raises the question of whether some factors render the processing route for plural processing either more lexical or more grammatical.

A second objective of the experiments in this study is to find the factors influencing the nature of plural processing. The literature reviewed in Chapter 5 will show that previous work on inflectional processing has supported the idea that various semantic as well as formal factors (e.g. the regular vs irregular distinction, typological differences in inflection, frequency) influence the processing nature of inflections.

The third question addressed in this project is the following: what happens in bilingual processing of plural inflections? In particular, at what point do the above-mentioned factors influencing the processing route emerge in different learner groups? This question is of major importance to the field of second language acquisition. If the factors influencing processing paths could be shown to appear at a given point in the acquisition process, educators could make a clear statement about the learner´s proficiency in terms of processing and *learning,* as described by Kupferman in section 1.1, could be made visible. If, however, these factors should not show up in processing although the learners use the plural inflection correctly, a qualitatively different processing route for learners could be assumed.

1.2 Outline of the thesis in summary

The following chapters will give a more detailed overview of the theoretical issues in the field and will place this study in a theoretical framework. In Chapter 2 lexical organisation will be described. In a thesis about inflectional processing claiming some intersection between lexicon and grammar, lexicon and grammar should be defined and explained in terms of the linguistic mental model. Consequently, Chapter 3 introduces the most influential theoretical models that attempt to explain the nature of access to the lexicon described in Chapter 2. Based on the static description in Chapter 2, of how language representation is organised mentally, Chapter 3 introduces some models of linguistic processing within such a hierarchical organisation. This is of interest because the experiments in this study deal with single words or noun phrases and their processing in real time. Only if we have a model of lexical access will such reaction times actually make sense and will it be possible to make claims about the way a speaker mentally accesses a stored word in the mental lexicon. Reaction times will not tell us anything without a model to interpret them. Chapter 4 then gives a brief overview of the purpose of the morphological structure in question, the plural. Chapter 4 is mainly a descriptive account of how the plural is being used in real language since it is not merely an arbitrary structure with formal properties. Moreover, since the plural has a purpose in language and some languages do not even encode it morphosyntactically, it is necessary to consider its function in language. Chapter 5 returns to the formal properties of plural processing. After having introduced the plural and its function, this chapter provides an overview of the myriad studies related to plural processing. These studies provide the background for the development of the two experiments introduced in the subsequent chapters. Chapter 6 explains the two experiments used for this study, a lexical decision task and a phrasal grammatical judgment task. Chapter 7 summarises the main results of the two experiments. A general discussion as well as some general conclusions will follow in Chapter 8.

2 *Lexical Organisation*

Chapter 1 gave a brief introduction and focussed on the aim of this study. Chapter 2 will provide the necessary theoretical background by introducing the psycholinguistic model of language representation which serves as the basis for this study when reference is made to lexicon and grammar as two separate components in language organisation. The seeming independence of the two components will be discussed in terms of their formal and functional properties and in their interrelation.

2.1 Meaning and form

Psycholinguistic theories often distinguish between various components in the language production process. Components are often thought to be modular, i.e. they function as subparts which operate more or less autonomously with a very specific type of input only. According to Levelt´s model of language production, such subcomponents can operate in a parallel manner because they work on just one particular part of the complete production process (Levelt, 1989). These modules are understood to interact only through interface representations, i.e. on the level of their input and output. This view automatically implies that the lexicon is considered a separate module containing all the lexical entries of the speaker´s language[4] and their semantic characteristics. These entries contain a specification of the syntactic function and the syntactic environment in which the lexical entry may appear (Levelt, 1989; Janssen, 1999). As such, lexical entries to a large extent already contain grammatical information. This is supported by speech error research and the tip-of-the-tongue (TOT) phenomenon (Levelt, 1989; Shattuck-Hufnagel, 1983). Speakers in a TOT state can correctly guess the phonological, lexical and grammatical properties of a word without being able to name it. These studies show that conceptual, lexical and grammatical information is available without the complete phonological form realisation (Vigliocco et al., 1999; Bowers et al, 1999). Syntax, therefore, is neither conceptual nor purely post-lexical but part of the syntactic information tagged onto the abstraction of a word, the lemma which is contained in the lexicon. This was also demonstrated in research by Vigliocco et al (1999) and Bowers et al (1999) who studied TOT phenomena as well as an experimental picture and word categorisation task for grammatical gender as well as for the mass/count distinction of nouns. Both gender and mass/count information are accessed faster in the word categorisation task than in the picture categorisation task, suggesting that such grammatical decisions operate on a lexical level. Also in the TOT state, subjects were able to correctly guess these syntactic features such as the mass/count and the gender distinction. Such studies can be seen as evidence for a separate representation level between conceptual representation and word-form representation in language processing. Such a separate lemma level has also been accounted for in the Levelt (1989) model.

2.2 Levelt´s model of language production

The aim of this study is to describe the processing path of inflected words and determine whether they are really purely grammatical and processed in a unit separate from the lexicon. In order to study the extra-lexical generation of inflectional morphemes in language processing, one has to look at the nature of the mental lexicon. It is not self-evident to speak of the lexicon as a separate functional entity. In fact, recent neurological studies have demonstrated that lexical processing may depend on exactly the same neurological structures as other semantic processing. In this case, "...there is no need to propose a *lexicon* in the brain - lexical items are not located in a generalized lexicon

4. Whether it contains only the entries of one or the entries of several languages is being debated (cf section 2.2.2)

module, but rather are located in the regions involved in the processing of the semantic features of the word or to the type of input (auditory, visual, or motor)" (Jones, 2004: p.112). Jones proposes two neurological structures that are distinguished in the role they play in memory consolidation; she distinguishes episodic from semantic memory and views these general cognitive modules as the neurological correlates of the more specialised linguistic functions. However, ample empirical evidence (briefly sketched above and summarised in Levelt, 1989) supports the suggestion that the lexicon could be seen as a functionally autonomous processing unit. For the case of simplicity and because of the vast evidence in favour of a separate functional entity, the existance of a lexical processing unit will be assumed in this study. This does not exclude the possibility of a vast amount of overlap between the specialised neurological structures for different processing tasks. Rather, the separate modules should here be considered symbolic. It is hard to specify the exact nature of linguistic processing modules such as the lexicon. Jones (2004) denies the existence of a separate lexicon and proposes a modularity based on different kinds of mnemonic processing. In this work, it is irrelevant whether the lexicon is a mnemonic unit or a purely linguistic one. It is easy to assume that it has some functional autonomy. And because of the neurologic evidence we have today, it is hard to sustain that the neurological structures of "the lexicon" are exclusively used for just lexical processing. Instead, there is a lot of neurological overlap for different functions. Nevertheless, it is still safe to assume that some functions have a certain processing autonomy and can hence be considered a symbolic module for the sake of simplicity. In Levelt´s model of speech production (cf. figure 1), the lexicon plays a central role in the speech production process. Levelt acknowledges that syntactic processing is lexically driven (Levelt, 1993) and therefore steers all other mechanisms. The role of the lexicon is central for this study since the lexicon contains declarative, abstract knowledge about lexical units, as described above. Levelt distinguishes two stages in lexical processing. He speaks of lemmas and lexemes as separable units in lexical encoding, both part of the lexicon. Whereas lemmas merely form "the intermediate level in the connection from a conceptual specification to a word form" (Janssen, 1999: p.17) and contain the semantic notion of the word together with information about the syntactic properties, the lexeme already carries word form information as well as phonological information. On the lemma level, these form specifications are not yet developed.

Levelt´s model, also called the "blueprint for the speaker" (Levelt, 1989) consists of several processing components that share a number of characteristics, mainly that they are relatively autonomous and independent in their functioning (Levelt, 1993). This also becomes apparant in Figure 1 where both the lexicon and the formulator are depicted as separate components. The lexicon is clearly divided into a lemma and lexeme level, the latter coming into play after an information exchange with the formulator has taken place for spelling out the form specifications.

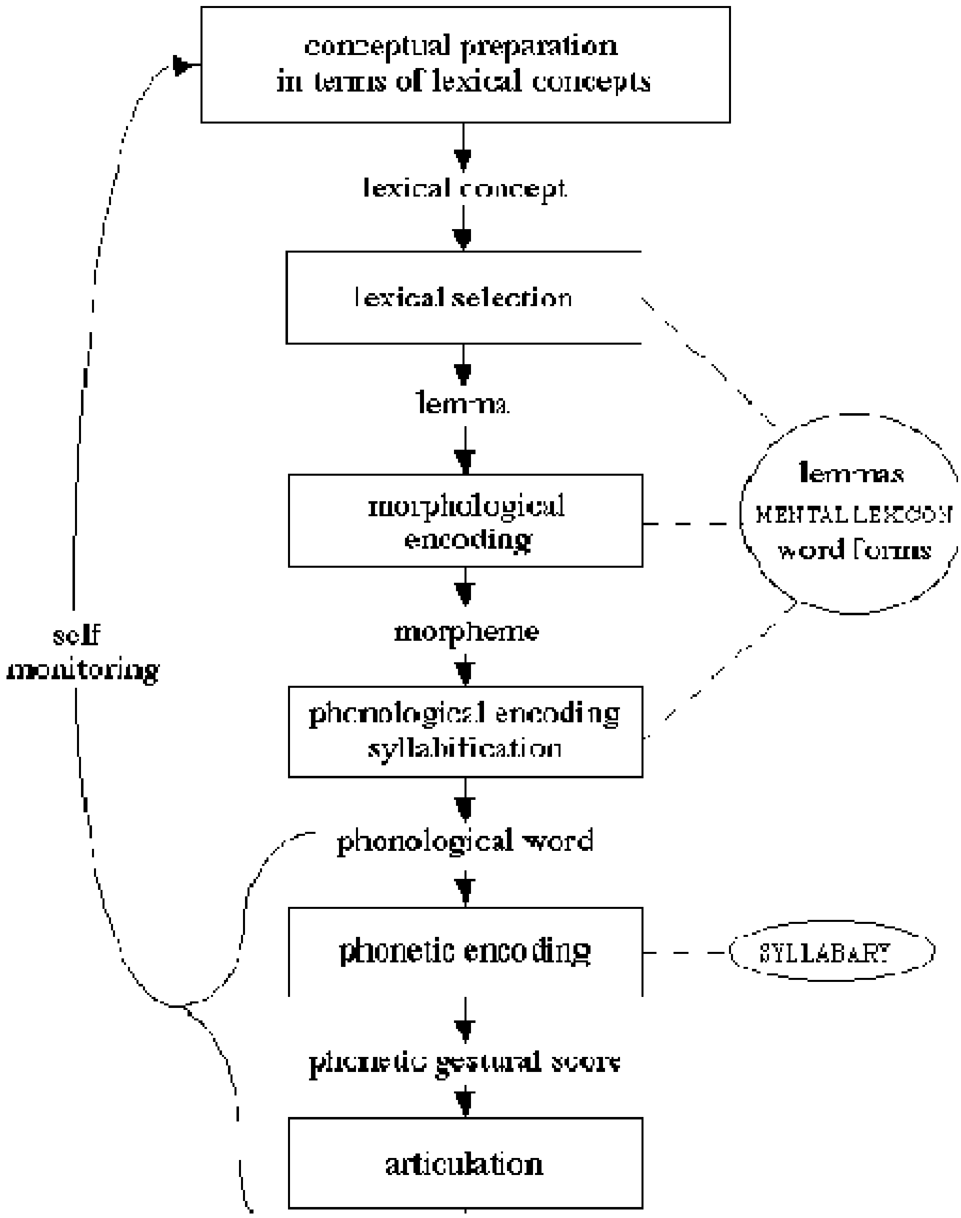


Figure 1. *Levelt´s model of speech production: boxes represent processing components; circle and ellipse represent knowlege stores (Levelt, 1993)*

2.3 Modularity

The modular nature of the speech production and comprehension process, as sketched in Levelt´s model, automatically poses the question of how the separate processing components interact with each other. At first sight, such a modularised model suggests a clear division between meaning and form, which is also supported by neurolinguistic research (e.g. Ullman, 2004; Vigliocco, 2000). A modularised theory of language processing differentiates between lexicon and grammar, as can be clearly seen in Figure 1. This distinction between meaningful and form-based processing is mirrored in morphological processing (cf. section 5.1) where meaningful and form-based morphemes are assumed to have different processing momentums. Derivational morphemes are, thus, thought to be processed in the lexicon and inflectional morphemes in the grammaticaliser (cf Perlmutter, 1988; Badecker & Caramazza, 1989; Schriefers, Friederici, & Graetz, 1992, Pinker, 1999). This assumption is known as the split morphology hypothesis because morphemes are split into two subgroups with different processing routes.

A clear functional division between a stored knowledge component for rote learning and a processing component for procedural operations in the brain has also been posited by Ullman (2004), who associates these polar functions with two distinct brain memory systems. Ullman provides evidence from neuroimaging studies as well as from the linguistic behaviour of patients with language disorders and with non-language disorders. All the linguistic behaviour observed in these cases points to two fundamentally different brain sytems in terms of localisation and function: the procedural memory system and the declarative system. Moreover, Ullman argues that these brain memory systems are not language-specific but serve other non-language functions in both humans and other primates. Ullman (2004) therefore claims that

> The basic premise of the DP (declarative/procedural) model is that important aspects of the distinction between the mental lexicon and the mental grammar in language are tied to the distinction between declarative and procedural memory.

The consequence of this assumption is to deny that language is a unique cognitive system. Not only are these brain areas not language specific, they also are not clearly separable in terms of domain specificity. Such a view has also been advocated by Jones (2004) who argues for one neurological learning system, independent of the nature of the task involved.

The learning system, she argues, treats lexical learning just like any other kind of learning, be it episodic, spatial or motoric. A brain area that underlies the procedural or the declarative memory system may partly serve other cognitive functions. Thus, the brain components involved in either system are not domain specific and have general cognitive functions. Only in combination and in interconnection with other brain areas do they serve specific functions. Ullman (2004) also adds that "it is not expected that all parts of the brain system underlying procedural memory subserve all aspects of the mental grammar, or that all parts of the brain system underlying declarative memory subserve all aspects of the mental lexicon". Similarly, he acknowledges that not all aspects of language can be reduced to the two systems (Ullman, 2004). Ullman´s position is central for the discussion about the processing distinction between lexicon and grammar posited in this work. However, also Ullman (2004) acknowledges a certain degree of autonomy of the mental lexicon and the mental grammar (cf. above quote). Jones (2004) is more radical in claiming that none of the brain areas involved in memory consolidation for any kind of learning task are specific; instead they serve all kinds of learning functions. Lexical learning, in her view, can be characterised as memory consolidation of the different lexical properties such as meaning, grammatical classification, pronunciation or information about animacy, colour or motor associations. Lexical processing is therefore based on a "dynamic, integrated neural network" (Jones, 2004: p.124) and not at all restricted to a lexical module.

Modularity is the key issue of this thesis. The view adopted in this work is a modular one, albeit with a very limited modularity. The functional distinction between grammar and lexicon has shown to exist. The fact that neurological processing systems overlap and that parts of the brain are not functionally specific does not mean that a symbolic modularity cannot be assumed. I will provide sufficient support in favour of a dinstinction between lexicon and grammar for morphological processing in particular in Chapter 5. The purpose of this work is to define the nature of the modules and the "degree of modularity". In order to understand a modular language processing theory as illustrated in Levelt´s (1993) model (cf. Figure 1) about language production, the separate components and especially their interaction will briefly be explained in the following subsections. It is important to bear in mind that the lexicon is the central component in this process (Levelt, 1989) and all interaction with other components will be considered under this premise.

2.3.1 The nature of words

The lexicon is the storehouse for abstractions of words. When speaking of abstractions of words, of lemmas, and of lexical information, it is relevant to consider the nature of words. It is clear that lemmas should contain some semantic information. After all, it is the conceptual thought that calls upon the lexical information in the lemma. This semantic information, however, is not enough to define a lemma. There is no one to one relationship between meaning and word and accordingly, none between meaning and lemma. It is perfectly possible to express a single idea in several words, or vice versa. It is, for instance, possible to say that something is „out of order“, which neatly describes a single idea in three words, whereas the single word „approach“ already expresses two ideas, namely that something is moving and that it is coming closer at the same time. The meaning of a word does not suffice to define a word as a single unit.

Singleton (1999) explores the possibility of defining a word by means of its orthographic, its phonological, its morphological or syntactic characteristics. All of these possible definitions, however, seem problematic, especially in the light of enormous typological discrepancies. Words are best defined as structural units (Singleton, 1999). If words are to be seen as structural units, this certainly holds for lemmas, as well. Such a structural definition already presupposes that meaning is not the core of the lemma; rather a combination of meaning and several structural properties make up a word. This is also in keeping with the assumption that much syntactic information is already present in the lexicon, albeit not yet specified (cf. section 2.1). Structure, in this case, cannot be viewed as a distinct property. This view poses some serious problems for the modularity assumed before. If the core of words is structural, some structural information is by definition part of the lexicon. Levelt (1993) has acknowledged that the lemma contains a structural frame. As such, it is already an integrated part of the lexicon, even though the exact grammatical properties are not yet specified in the lexicon. But exactly this interplay between grammatical and lexical information is the topic of this work on morphological processing.

In conclusion it can be said that the lexicon bears psychophysiological real structural and content information that cannot be spearated, not even on the lemma level. This is the basic tenet of the present study, which concentrates on the emergence of native-like processing in language learners. This brief discussion about the nature of words will be followed by a discussion about the nature of the connections between the separate language processing components.

2.3.2 Lexicon and conceptualiser

Despite the importance of structural information for the essence of a word, it is nevertheless the meaning and its communicative function that induces a speaker to access the mental representation of a word. A mental representation of both meaning and structural information is assumed to be stored in the lexicon at the lemma level (Levelt, 1993). In lexical encoding, the two processes may be viewed as separate and sequential in nature, with the meaning access preceding lexico-syntactic encoding (Vigliocco et al., 2002). This is the case, at least, in on-line speech production. The lexicon can therefore be viewed as an autonomous module in terms of processing sequence, functionally separated from the conceptualiser and containing some grammatical information (Levelt, 1993; Bowers et al., 1999). Experimental evidence including the above-mentioned (cf. section 2.1) processing of count/mass and gender information in either picture or word stimuli, supports a model that functionally differentiates between lemma level and conceptual level (Bowers et al., 1999). An autonomous lemma level allows lemma-processing without any interference from the coneptual level, for instance.

Various models have described the relationship between words and concepts within a monolingual or bilingual speaker. Since the first influential description of linguistic representation in bilinguals by Weinreich (1953), these models have been slightly amended to match them with latest empirical findings on lexical representation in bilinguals. Weinreich distinguished between the co-ordinate, the sub-ordinate and the compound bilingual system. One scholar who extended the original models is Kroll (1993) who introduced the revised hierarchical model of lexical representation. In this model, connections between conceptualiser and lexicon are stronger in the first language than in the second language. In empirical studies, Kroll (1993) also found stronger lexical connections from the second to the first language than vice versa. After all, learning a second language implies that one tries to find lexical equivalents in the first language. Therefore, second language lemmas are closely associated with first language lemmas whereas a word in the first language does not automatically call upon its second language translation equivalent. Kroll (1993) distinguishes between two lexical stores, one for the first language and one for the second language, while assuming a shared conceptual storage unit for both languages.

The importance of distinguishing between a conceptual and a semantic level in language representation has repeatedly been proven (Levelt 1993; Kroll, 1993; Pavlenko, 1999); these experiments clearly reveal an experimental effect of “meaning” on a word level that is absent in non-linguistic (picture) tasks. Although researchers do not agree on the degree of semantic overlap between the two languages in bilinguals, several studies have shown evidence of a dynamic system where a multitude of factors may influence this degree of overlap between the languages, most notably language development (Pavlenko, 1999; Kroll, 1993; Kroll & Curley, 1988; Chun & Ho, 1986; Mägiste, 1984; de Groot, 1992, Talamas, Kroll & Dufour, 1999). De Groot (1992) claims that all three structural representations proposed by Weinreich (1953) can be found within an individual bilingual. She says that word class or status determines the model presently available for the speaker. De Groot summarizes these factors as word-type effects. After a series of experiments, De Groot and Nas (1991) concluded that bilinguals have a shared lexical representation for cognates, as opposed to noncognates. In Weinreich´s (1953) terms, learners at different developmental stages employ different bilingual systems when speaking. Weinreich did not acknowledge that his three types of bilingual representation could occur within the same person. Rather, he thought of different types of bilinguals depending on the nature of their second language acquisition. Nowadays, however, the question has shifted towards the strength of connections between the lexical items, thus recognizing mixed representational systems within one person. Therefore, the transition from one system to another can only be interpreted as a gradual development in terms of moving from one end of the scale to the other end, without completely changing the processing mechanisms. In a real dynamic system there is constant fluctuation rather than a linear development; acquisition and loss are both possible in development.

De Bot (1992), who extended Levelt´s (1989) model to bilingual processing, assumes that the conceptualizer already contains language specific information. He distinguishes between two stages in conceptual processing. The first is to be seen as language independent and the second as language specific. What is crucial in de Bot‘s assumption is that, in contrast to Kroll´s (1993) model, bilinguals have only one lexicon, regardless of the language. Even if one is to assume a unitary lexicon, lexical items need to be interconnected. Several researchers in the 70s and 80s (Fromkin, 1971; Laufer, 1989) claimed that the way we organise and process lexical knowledge when two or more languages are involved primarily reflects phonological connections between the lexical items, rather than semantic connections. Meara (1984) carried out a study on lexical interconnections and found evidence in support of this hypothesis. However, the results are ambiguous and can be interpreted in a different way (Singleton, 1999) that also supports the notion of semantic connectedness. Later studies have found evidence counter to the phonological assumption and recent psycholinguistic evidence contradicts the strict phonological view (Singleton, 1999). In addition, individual learners seem to shift as they become proficient in the second language. Early interlexical connections might be of a rather phonological nature, but the more the semantic system in the second language develops, the more lexical connections become "semantic". This is exactly what Hulstijn and Tangelder (1993) found when they tested the extent to which English word pairs similar in form and/or meaning were confused by learners of English at different levels of proficiency. This is in keeping with the shift in responses to word association tasks. Whereas early learners may prefer to give syntagmatic associations, more proficient learners tend to give paradigmatic associations to a given word (Singleton, 1999; Hakansson, personal communication April 2004). Jones (2004) even claims that accessing a single word from the lexicon involves the activation of a variety of word features, irrespective of the language in question. Likewise, she assumes that the activation of a word automatically entails the activation of the conceptual information, thus questioning modularity altogether. The same would be true for monolingual processing. However, she does not address the question of how a learner passes from a monolingual to a bilingual state. There seems to be no single organisational representation in the bilingual lexicon. The question is whether one has to assume different lexical organisations within one individual, depending on the situation and level of proficiency. This suggests that static models of language representation might not adequately describe the mental processes and might underestimate the role of developmental and probabilistic factors.

It seems clear that the co-ordinate system, suggesting complete separation between the two language systems in bilinguals (both lexically and conceptually) is difficult to maintain regarding the amount of interlanguage transfer, code-switching and the ability to translate between languages. Nevertheless, bilinguals must possess a means to „switch off" one of their languages when using the other language in a strictly monolingual context. On the other hand, bilinguals can always rely on their „sleeping" language even in completely monolingual tasks (Cenoz, 2002). It seems to be the situation rather than the acquisition process that determines which representational system is the working model within a bilingual. This is a nice parallel to Tuggy (1997) who stated that the situation and the linguistic context determine lexical versus grammatical processing (cf. 2.3.3). So even within the lexicon such situational, contextual and probabilistic issues seem to play a role. In terms of lexical access models, this has been put forward by Morton (1979) in his logogen model that acknowledges probabilistic factors in word retrieval. A unifying conclusion from all these studies may be the "fuzziness" of such abstract concepts as the language modules we are talking about. The idea that a multitude of factors influence the processing of language and the degree of "separability" of such modules is also the key line of thought pursued in this study. In sum, bilinguals may have different lexical representational systems at work. These depend on the bilingual's proficiency in the two languages, and the word types as well as the tasks carried out. Bilinguals may also have different processing strategies at work or use them to different extents and at different times in their language development. The factors influencing the first and second language processing route will be explored in this project where learners and native speakers are involved.

2.3.3 Lexicon and Grammar

Lexis and syntax can be viewed as entangled and mutually dependent. Even though some of this grammatical information in language organisation needs to be generated after lexical retrieval, a lot of grammatical information is already present in the lexicon (Vigliocco et al, 1999; Bowers et al, 1999). This study queries the precise dividing line between lexicon and grammar and suggests that both modules need to be defined in processing terms. Furthermore, it shows that the boundaries are fuzzy with respect to language development and language situation. Processing is influenced by a multitude of factors and cannot be simplified by referring to a clear modularity of functions. Although it is easy to separate the function of grammar and lexicon in order to explain language processing, it should become clear after reviewing the literature and the set of experiments in this work that the separation of the two modules should be considered symbolic. Although some of the research points to a strict functional separation between the modules, other research stresses the fuzziness of the modules. In agreement with Ullman (2004) I would suggest that the functional separability is psychologically real. However, its cognitive correlates do not exclusively serve a linguistic function. Although there may be clear boundaries between the different modules in some processing tasks, others may rely on a more dynamic processing system. As such, processing is very much task- and situation dependent and it is also very personal.

The focus of my experiments lies on the processing of the plural *"-s"* morpheme. In morphology, the assumption of a distinction between inflection and derivation presupposes a separate lexicon and grammar. As Tuggy pointed out in his 1997 paper, grammar has traditionally been viewed as rule governed, whereas the lexicon has been thought to contain idiosyncratic and rote learned items. He argues that this dichotomy should be rejected, since the distinction is not as polar as originally assumed. Tuggy (1997) claims that any word form or phrase form can be processed both ways, depending on speaker, situation and storage. As such, the human language processor permits redundancy. The same inflected noun may either be processed as grammatical in a linguistic context where it appears among other inflected forms or else it may be processed as lexical in a situation where it is learned and remembered as a single lexical item. Such a dual route processing mode has also been suggested by Baayen, Dijkstra and Schreuder (1997). The notion of a clearly discernable lexicon and grammar is therefore challenged. Also for morphological processing, the two components do not seem as separable as they might first appear.

For bilinguals, de Bot (1992) proposes a bilingual version of the Levelt (1989) model of language organsiation (cf. section 2.2) in speech production:

> ...There is one lexicon where lexical elements in different languages are stored together. It has been suggested that the relationship between the lemma and form characteristics in bilinguals is not one to one as in the unilingual case: a lemma can be linked to various form characteristics depending on the language or languages involved. Within the lemma, meaning and syntactic information may not be inextricably linked (de Bot, 1992: p.21).

De Bot does not specify the way in which lemmas and lexemes may be connected. Language relatedness in bilinguals as well as level of proficiency seem to play a role for the nature of the structural representation in a bilingual‘s mental lexicon. He does, however, suggest that meaning and syntactic information within the lemma are not as tightly interwoven in bilinguals as in monolinguals, which would make sense if one considers the seeming flexibility that bilinguals display, for instance, in code-switching, where form information of one language sometimes transfers into the other languge (Broersma & de Bot, 2001).

A bilingual model of language production has to account for two language systems that can be used entirely separately or mixed. It also has to allow for various degrees of bilingualism without assuming a loss of speed in speech production. Of course, it needs to cope with more than two languages and also languages that are typologically not so related. Another question is how strong the links actually are between meaning and syntactic information in the lemma. As a lot of syntactic information already exists in the lemma, an interesting point to be explored is whether this information is as strong in second language learners as it is in native speakers. The experiments described here should help to clarify the problems in defining a bilingual model of language organisation and they should attempt to solve the questions of the extent to which lexical and grammatical information is idiosyncratic and how this information changes in language development.

2.4 Language development

Bates and Goodman (1997) have taken a closer look at various typical and atypical learner groups and found that their psycholinguistic data supports a lexicalist theory of grammar. Like Schönefeld (2001), who gives a concise overview of evidence for the interrelationship between grammar and lexicon, Bates and Goodman (1997) supply ample evidence for the close association between the two.

First, they observed data from first language acquisition of English and compared the lexical and the grammatical development. For this purpose they compared the acquisition of vocabulary items with the mean length of utterance (MLU) in morphemes in order to have an objective measure of grammatical development. The children´s development supports a strong correlation between grammar and lexicon in language development. Since the relationship between the development of function words and grammatical morphemes could, however, be considered a tautological one, Bates and Goodman (1997) also looked at first language acquisition of Italian, a language with quite a different typological division of grammatical and lexical features. Inflected plurals in Italian, for instance, are not added as suffixes to the word stem, but may change the word stem instead. For instance the word *pizza* changes to *pizze* in the plural. Without the plural suffix, the remaining stem would not be a complete word. Even in the case of Italian, Bates and Goodman found a clear correlation between lexicon and grammar in language development.

The same findings hold true for children with atypical language development, i.e. very early and very late talkers. Despite differences in utterance length in the children examined, Bates and Goodman could not observe any dissociation between grammar and lexicon. Other children with atypical language development include those with unilateral brain damage. These children seem to show an overall delay in language development, in terms of both lexicon and grammar (Bates & Goodman, 1997). Moreover, lesions in neither the left or the right hemisphere, produced any of the aphasic effects observed in adults with brain injury. The specialised brain functions such as the Broca and Wernicke area in adults are not yet developed in children. This illustrates that grammatical and lexical abilities "are effected equally by early focal brain injury" (Bates & Goodman, 1997: p.14).

When Bates and Goodman studied the language development of children with Williams syndrome and Down Syndrome, it seemed that the children with Down syndrome showed the most dissociation between lexical and grammatical development, with the development of grammar lagging behind. Bates and Goodman (1997) nevertheless, speculate that this dissociation might be caused by atypical information processing only indirectly related to grammar, with impairments of auditory processing and short term memory, in particular. Reviewing language disorders in adults, Bates and Goodman likewise found no clear dissociation between lexicon and grammar. Rather, the non-fluent pattern of language in adults with language disorders impedes both lexicon and grammar and they break down together. Finally, Bates and Goodman (1997) propose that the syntactic breakdown may rather be a morphological breakdown because of the non-salient nature of closed-class morphemes in language perception. Therefore these morphemes are prone to be deleted when the processor is under stress. This has, in fact, been demonstrated in a study with healthy and linguistically normal college students. The processor under stress appears to be a crucial factor in agrammatic language production. Since such a processing breakdown may occur in healthy and impaired populations, it is not caused by a specific localizable area in the brain. Bates and Goodman (1997) conclude that there is very little evidence for a dissociation between lexicon and grammar and they claim that, so far, the evidence for a pathological dissociation has not been convincing. Again, this research supports the assumption that there are no strict boundaries between the functional language modules in language development. Another reason to assume that processing is a dynamic event, its nature depending on the stage of development and relying on general cognitive functions.

In contrast, Ullman (2004) functionally differentiates two distinct brain memory systems which he associates, if not on a one-to-one basis, with grammatical and lexical knowledge in language. Although he finds evidence for such a functional distinction in the human brain, he also states that these systems are not unitary and domain specific but they overlap. This means that local parts of one functional memory system in the brain which are closely related with grammatical and language independent route processing may be equally involved in the other functional memory system which is related to rote learning and lexical knowledge. Sub-components can always serve various cognitive functions. For learning tasks, Ullman (2004) even presents experimental evidence that suggests that learning one and the same cognitive function (e.g. grammatical rules) may depend on both functional memory systems, depending on the point of development. In early learning, the declarative system may play a stronger role and later in the learning process there seems to be a shift towards procedural processing. Here may be the link between the neurological findings and the two distinct learning systems of explicit (declarative) and implicit (procedural) learning as proposed by, for instance, Krashen (1983) a couple of years earlier. Here, the basic question is whether the learners shift from one processing mode to another during language development and to what extent they approach the native speakers´ processing behaviour. For single words, either uninflected or inflected, different strategies may be at work.

Summing up, with the various findings on language representation advocating a modular view (Ullman, 2004; Vigliocco, 2000; Vigliocco et al, 1999; Bowers et al, 1999; Levelt, 1993), it seems impossible to deny a certain degree of functional modularity in language processing. It is, however, not clear whether this functional modules may be due to a general cognitive distinction between procedural and declarative memory (Ullman, 2004) or whether there may be some purely linguistic modularity. Despite the evidence for a modularity in some tasks of some speakers, a large group of researchers (Singleton, 1999; Pavlenko, 1999, Talamas, Kroll & Dufour, 1999; Kroll, 1993, de Groot, 1992; Kroll & Curley, 1988; Chun & Ho, 1986; Mägiste, 1984) propose a more dynamic system of language representation where the nature of processing depends on the stage of language development in the first or second language, the situational context and the linguistic task. Speakers may rely on different functional memory systems when processing one and the same linguistic structure, depending on a multitude of factors. This will also be shown in the experiments reported in the later chapters. I therefore support a limited modularity in language processing, in which the linguistic modules are symbolic for functional brain areas that can operate exclusively for a specific linguistic function or partly for a multitude of other cognitive functions. This is clearly a view that does not understand processing in a completely interactive manner without any functional separation as suggested by Jones, 2004. But it is a model that understands language processing as a multi-dimensional event which may take several paths depending on the speaker´s situation.

.

3 *Lexical access*

The previous chapter introduced the model of language representation that forms the basis for this study. The module components of this model have been introduced and their interactive role has been explained. This chapter concentrates on the role of the lexicon and on the processes involved when a lemma needs to be retrieved from the lexicon. Although it is problematic to separate lexical design and lexical access since they are functionally related, most access theories do not address the organisation of the language faculty. Lexical access will therefore be treated as a separate issue in this chapter. It will explore the main theories of simple lexical access in language recognition. This overview is needed for a deeper understanding of the morphological theories as well as the on-line experiments reported in the chapters following. The most commonly known theories of lexical access in language recognition will be discussed in the light of the discussion about the nature of inflectional encoding.

The process of accessing words from the mental lexicon, is subject to the nature of the word in question. Inflected words are generally supposed to rely on other processing mechanisms than uninflected words and inflectional morphology behaves differently from derivational morphology (cf. 5.1). This assumption also follows from the Levelt model presented in Figure 1 and will be discussed in various morphological processing theories in Chapter 5. In order to define the processing paths for different kinds of words or different kinds of morphemes, it is necessary to define the types of morphemes themselves. Several theories distinguish between either theoretically or functionally different morphemes.

Pienemann´s Processability Theory (1998), for instance, predicts that learners acquire certain morphosyntactic structures of a language in a strict developmental order, depending on the kind of information exchange between the phrasal constituents. The sequential distinction betwee the mophemes in acquisition, thus, results from processing differences. Therefore, plural inflection on single words without a grammatical context and as a single constituent (e.g. *cats,* as a single word in a context where one points to a group of cats) occurs at an early stage in language development when conceptual structure may be directly mapped onto linguistic structure. Plural agreement inflection involving two constituents (e.g. *two cats*, determiner and noun in agreement) occurs at a later stage in language development because, here, the two constituents need to exchange grammatical information. Whereas in the first case, the grammatical information (plural) is stored in the lexicon, encoded as part of the concpetual information, it has to be exchanged in the second case (plural agreement). Pienemann (1998) proposes that it is the functional nature of the inflection that determines its processing route. He argues that one and the same morpheme, in this case the plural-*s* inflection, may employ two different processing procedures depending on the linguistic context. Processing of morphemes is therefore highly dependent on linguistic context.

A similar idea has also been put forward by Booij (1993) who argues that there are two types of morphemes, inherent and contextual ones. Whereas the former are a necessary addition to the word stem in order to render the word meaningful in the specific situational context, the existence of the latter is based on grammatical context only. Booij (1993) therefore makes a distinction between two inflectional types based on the function of the inflection. This view reflects a modular processing model for language production that clearly distinguishes between the lexical and the grammatical component of language as discussed in Chapter 2. This means that lexical access and grammatical encoding are considered to take place in two functionally different linguistic modules: direct lexical access from the lexicon and subsequent grammatical encoding in the grammaticalizer. The existence of these functionally and temporary distinct processing modules may be mirrored in the course of linguistic encoding. Since grammatical encoding is thought to take place after lexical access[5], this time difference should be explicit in online linguistic research employing reaction time techniques. The power of reaction time (RT) differences between different experimental conditions in online techniques is considered to reflect different temporal and possibly also locational processing in the human mind. Even though the nature of inflectional processing is often considered to be purely grammatical and expected to take place after lexical encoding, it may not be so uniform and it may -as Piemeann (1998) and Booij (1993) have acknowledged- depend on contextual, conceptual and functional factors, as well.

3.1 Comprehension and Production

Reaction time experiments can tell us something about the processing quality of inflected nouns. Some affixes may be part of a lexical entry; others may have to be seen as components adjoined to the lexical entry by means of grammatical rule application. The debate about the nature of inflections and, therefore, plural inflected nouns will be approached as a debate about lexical and grammatical processing in language production and in language comprehension.

It should be mentioned that the Levelt model (cf. Chapter 2) was primarily developed for speech production although later models included comprehension. The experiments in this study relate only to language comprehension. It cannot automatically be assumed that the processes involved in comprehension are the same in a reverse order as in language production. Nevertheless, the issues addressed here are central to both production and comprehension. Any processing model must distinguish whether the processing nature for lexical access -when either searching a word before uttering it or when searching an acoustic stimulus in order to comprehend its meaning- allows a full storage of inflected nouns and verbs in the lexicon or whether the mental lexicon contains uninflected word stems only. In language production, separate processing of inflections would entail a grammatical encoding of the word stem after it has been accessed from the lexicon. In language comprehension, separate processing entails stripping off the inflection from the

5. cf the Levelt model in Chapter 2, Figure 1

word in order to access the word stem in the mental lexicon for meaningful comprehension. In both cases, a separate encoding of stem and inflection leads to a longer process than in accessing the full inflected form stored as a whole form in the lexicon. Thus, these issues are the same, regardless of the direction of processing.

The theories of lexical access introduced in the following sections will refer to lexical access in word recognition. This is also the task the subjects face in the lexical decision experiment described in Chapter 6. The following lexical access models will be discussed especially in the light of the explanations they provide for inflectional processing.

3.2 Lexical access models

Levelt´s (1993) model of language production as briefly introduced in section 2.2 subdivides the mental lexicon into two separate levels, each comprising specific functions in the speech production process. The existence of these two levels has repeatedly been proven through experimental psycholinguistic evidence as summarised in Levelt (1989). This research has demonstrated that the first level, the lemma level, can clearly be discerned from the second, the lexeme level, both in terms of access time and access nature (e.g. Levelt, 1989; Vigliocco et al 2002; Bowers et al, 1999).

When retrieving a word in speech production and when accessing the lexicon in language comprehension, the language user needs to find the appropriate lexical item in the vast amount of entries in the lexicon. This seems an amazing endeavour when one considers the speed of lexical access. A speaker may utter up to six words per second without consciously attending to the processes involved. Thus, such a highly automatised and fast skill must follow some systematic procedures. Word recognition or word retrieval has been hypothesized to proceed in various ways. Some of these models of lexical access that explain inflectional processing will be looked at in more detail in the following sections. These models have to acount for the word recognition of both monolingual and bilingual speakers as well as for language learners (Garman, 1990; de Bot, 1992).

3.2.1 Search model

One much-discussed word recognition model is Forster´s search model for lexical access (1976). The model is a two-stage model, containing an access file - an operator file as it were- and a master file. The model functions like a search in a library catalogue: an entry is looked for on the basis of its form in an alphabetical order and the output given is an abstract location marker which tells us where to find the entry. The access file looks for an entry sequentially, letter by letter or phoneme by phoneme and can be compared to a library catalogue search (Garman, 1990). As a library catalogue search may be based on a variety of information available (name of the author, title or year of publication), lexical access may depend on different form input, either phonological, orthographic or semantic. Consequently, there are various access files, depending on the mode of word representation but only one storage system, the master file (Garman, 1990). All these form files, however, eventually lead to the same entry in the master file, just as all the catalogue searches would eventually lead to the same shelf in the library where the book can be found. Unlike the library search, the form files bear on psychologically plausible search mechanisms. The frequency of sound or letter combinations, for instance, may play an important role for the success of a search in the access file. The first search of the access file needs to be completed in order to access the entry on the basis of this abstract form representation on the master file. The actual entry then provides all other lexical information once it is found in the master file.

A model like Forster´s (1976) would predict shorter reaction times for identifying real words than for identifying nonce words. This would be the case because the access file would have to be searched completely for an infrequent nonce word, which in the end does not seem to be part of the access file. However, this conclusion can only be drawn after a complete search of the access file which is more time-consuming than a search that is interrupted once the first existing sound or letter combinations of a real word are identified. This assumption has repeatedly been criticised. The time-consuming refusal of nonce words as opposed to the relatively fast acceptance of real words is, however, also supported by the outcomes of the lexical decision experiment in this study (cf. section 7.1.4).

Forster´s (1976) search model also predicts longer reaction times for longer words as the search would have to go on for more items than in the case of short words. However, this prediction seems not fully compatible with the findings of various lexical decision experiments (Garman, 1990). Forster´s search model (1976) does account for the frequency effect found in such experiments, which suggests a frequency-based search of the access files with highly frequent forms being at the "top of the list". However, it does not give a satisfactory account for any semantic priming effects as observed in various other lexical decision experiments (Garman, 1990; Singleton, 1999). This frequency effect can also be found in the study discussed in Chapter 5 and generally seems to be a consistent finding in processing experiments (Garman, 1990; Levelt, 1989; Baayen, 2003; Bod et al, 2003). Garman (1990) and Singleton (1999) point out that some of Forster´s own findings seem contradictory to his proposed model. Although it is convincing in most respects, Forster´s model (1976) has continuously been subject to discussion.

3.2.2 Affix-stripping model

Another model that describes the time-course of lexical access is the model proposed by Taft and Forster (1975). Their model of written word recognition strips affixes from the word stem before it searches for the stem only, independently from the word´s possible affixes. Such a model predicts more processing effort for affixed words than for free lexical morphemes without affixes. Although it has been set out as a working model for prefixes, the model could be applied for suffixes, as well. In experiments Taft (1981) found longer reaction times for pseudo-prefixed words like "premium" than for non-affixed words like "lemon", but these experiments have been criticized for the stimulus material (Smith et al, 1984). Still, this model, with its hypothesis about separate affix processing, should be considered as one possible explanation for the findings of the inflectional processing experiments explained in Chapter 7. The affix-stripping model allows for economical storage (Garman, 1990) since only stems need to be represented in the master file while affixes are generated on the basis of derivational word-building processes as well as on the basis of grammatical rule application for inflections. This model reflects the traditional distinction between meaning and form, with all affixes and, therefore also grammatical inflections, being generated post-lexically.

3.2.3 Logogen model

A more general model accounting for word recognition is the logogen model by Morton (1979). Morton´s model was developed as a detector model for word recognition. The logogen model accounts for the probability of a word appearing in a given context, which envisages that some of the lexico-semantic and syntactic properties within the word will not be found in the primary search of the access file in Forster´s model, for instance. In Forster´s model, only formal properties are relevant for the first search, not probabilistic contextual information. Morton considers the lexical event to be a neural unit and coins it "logogen". In his model, input information needed for a lexical unit to be recognised may come from various modalities simultaneously. Sensory and contextual input may interact (Garman, 1990). Therefore, the general cognitive system is closely adjoined to the logogen system and, to some extent, to part of the lexicon, since it has to provide the knowledge of linguistic and extra-linguistic contexts in which a word may appear. Each logogen has a threshold that needs to be lowered in order to be "fired" (Morton, 1979).

The frequency of a word may influence its threshold activation. Frequency here is the frequency of occurence of a word in speech input, as well as its frequency within a given linguistic and pragmatic context. Thus, the logogen model is sensitive to frequency effects and it therefore accounts for the experimental frequency effects found in various experiments (Garman, 1990, and Chapter 7 of this study). High-frequency words have a lower activation threshold than low-frequency words and a specific linguistic context for a particular word may lower the activation threshold of its logogen. This is the case for collocation-like idioms with a probabilistically highly frequent linguistic environment, for instance.

Such a probabilistic frequency effect is also supported by Schönefeld (2001) who carries out a lexical decision task (described in Chapter 6) with repetition effects for complete phrases. That is, the subjects have to decide on the complete phrase, i.e. not on single words. She wants to find out whether the complete repetition of a collocation is processed differently from the repetition effects of the sum of the single words taken out of their collocational linguistic context. She compares these two collocation conditions with the repetition effects for complete free phrases and for single words taken out of free phrases. Schönefeld finds different processing paths for collocations and free phrases.

Interestingly, she does not find a significant difference between the repetition effect in exactly repeated collocations and in exactly repeated free phrases. Here it seems that both types of phrases seem to be processed similarly. This means that it does not make a difference whether a collocation or a free phrase is repeated, both seem to evoke a repetition effect. In both cases, the phrases are processed as an entirety, not as a sum of their component parts. It does not matter whether it is a collocation or a free phrase since both types have a semantic unity which allows them to be processed as a unit. Schönefelt does, however, find significant differences between the repetition effect in distributively repeated collocations (repetition effect is significant) and distributively repeated free phrases (repetition effect is not significant). In this case distributively repeated means that the component words of the collocations and of the free phrases are repeated as single words in a separate linguistic context. Schönefeld (2001) explains her findings by proposing that the occurrence of each single word -even in the distributively repeated condition- triggers the activation of the complete collocation, while these words do not trigger the complete free phrase which might be heard or seen before. Thus, the processing of collocations relies on different mechanisms from the processing of free phrases, even in on-line lexical decision tasks (Schönefeld, 2001). Crucially, Schönefeld provides evidence for the fact that the probabilistic co-occurrence of linguistic information is assumed to be part of a word´s lexical entry in the mental lexicon. So even one word in a free context triggers possible other contexts when activated. Schönefeld´s findings, therefore, support Morton´s logogen theory, which is based on the assumption that the logogen contains much more information than the mere formal properties of a word. The experiments on plural processing in this study will provide more evidence for the importance of probabilistic information in the language processor.

In this context it is worth mentioning that recently, Halgren et al (2005) have found neurophysiological correlates for word repetition effects in implicit priming and overt word repetition tasks. Halgren et al (2005) observed different activation effects in studying the behaviour of subjects with in situ electrodes while performing a word repetition task. The observations suggest that two functionally different neurological systems are at work depending on the nature of word repetition. Word repetition may be either accidental or deliberate, such as in tasks where repetition is task-relevant. Different structures are, thus, responsible for more implicit (e.g. *deja vu* effect) and more explicit recognition processes.

These structures, however, interact and are all part of the chained “activation cycles in an extended neural network that involves each of many cortical columns over sustained and cognitively varied processing stages” (Halgren et al, 2005). These observations can also explain Schönefeld´s (2001) different results on the repetition effects of collocations and free phrases, the former corresponding to the more implicit recognition described above. The observations by Halgren et al (2005) provide some very recent neuroanatomical evidence for the complex cognitive processes involved in word recognition as already suggested by Morton in 1979.

3.2.4 Cohort model

Marslen-Wilson and Welsh (1978) have proposed a model that considers lexical access as a search by auditory word detectors with gradually decreasing possibilities of target items until only one item remains and is chosen. As a person hears a word, the initial 150 ms of the acoustic input will be used to activate a cohort of all possible lexical entries with exactly the same beginning. All these candidates will monitor the subsequent acoustic information input, which will constrain the possible candidates until only one candidate will remain. Unlike the logogen model which sums up the partial activations until a threshold is being reached, the cohort model does not allow for only partial activation on the way to the optimal candidate (Garman, 1990) but fully activates all the possible candidates until a final decision has been made. Since the initial information is crucial for the first cohort of possible candidates, a nonce word with an illegal initial syllable would be spotted as a nonce word very early on in a lexical decision task; this is clearly the case in the lexical decision experiment in this study. As such, every word has a uniqueness point, a determined point at which the incoming sound signal clearly constrains the possible cohort of candidates so that only one lexical entry remains a probable candidate. This uniqueness point of word recognition has also been confirmed by experimental evidence (Garman, 1990). The model does explain context probabilistic factors in the recognition process. It accounts for such probabilistic information by assuming that it is already inherent in each cohort candidate. It does not, however, account for discourse-pragmatic contexts like the linguistic environment of the auditory input itself (Singleton, 1999). Therefore, experiments on linguistic context effects, such as those by Schönefeld (2001), cannot be explained with this model. However, this model does not make any assumptions about visual word recognition, which is relevant in the experimental tasks described above and also in the study to follow.

3.3 Implications for word recognition

Some of the models propose various ways of accessing a lexical item, depending on the mode of input. This distinction between various types of input mode must be psychologically plausible for the processes involved since the human being relies on different sensory information to identify a lexical unit. A spoken word may be perceived in a completely different manner from a written word. Spoken words usually do not have clear word boundaries and contain super-segmental information such as intonation and rhythm. Such information provides clues for the retrieval of the word in question during a continuous speech string. In written word recognition, however, the reader immediately recognises word boundaries and sees a word as a complete and discernable unit. To a certain extent, it is even possible to comprehend a word when the letters are exchanged within the word boundaries (Rawlinson, 1978; Perea & Lupker, 2003). So within the word boundary and within the frame of the exterior letters, jumbling of letters may not seriously hinder word recognition. Apparantly, different modes of language input have different implications for lexical access. Mode of input must consequently be an important factor for the access path.

Likewise, the difference in processing between real and nonce words seems psychologically plausible and is also supported in the lexical decision experiments described in section 7.1.4. In this lexical decision task, real words are recognised much faster than nonce words. The Cohort model (Marslen-Wilson & Welsh, 1978) would predict the exact opposite but Foster´s seach model seems to account for this phenomenon. These two lexical access models primarily focus on spoken language. Therefore, the outcome of these experiments may not have any implications for the theoretical validity of such models.

Also probabilistic contextual factors seem to have an influence on lexical retrieval. As discussed in section 3.2.3, Schönefeld (2001) found evidence for the relevance of contextual linguistic factors in language processing. Frequency will be shown to be a relevant factor in the experiments to follow in this work.

4 Number

The previous chapter provided an overview of the theories that explain lexical access from the lexicon in language processing. These theories have particularly been used to deliver explanations for lexical access of inflected words and for the processing factors that might be involved in lexical access. Since the nature of the plural inflection is the object of study and will be experimentally observed, it is necessary to provide some background information on the functional nature of the grammatical plural and on what the plural denotes. This chapter summarises several studies that scrutinised morphological processing related to plural inflections.

There are various ways of marking number in a nominal phrase. Obviously, one way of is to use overt lexical numeralizers, such as *one*, *two*, *three*, *many*, etc. In the English language, the plural (i.e. all number properties denoting more than one) in nouns also needs to be marked grammatically. In most cases the English plural is grammatically expressed by means of the suffix "*s*" which is therefore also considered the regular plural marking in English nouns. The various regular spoken realizations of plural markers are summarised in Table 1.

Table 1. *regular plural forms*

plural types	*plural form*
regular I (voiceless)	*cats, flats*
regular II (voiced)	*dogs, birds*
regular III (added vowel)	*houses, busses*

All of the plural realisations are to be considered allomorphs of the ‚s‘-plural. In contrast, a number of irregular plural forms exist, as summarised in Table 2.

Table 2. *irregular plural forms*

plural types	*plural forms*
irregular I (different suffix)	*oxen, children*
irregular II (change of stem)	*mice, women*
irregular III (converted from singular form)	*deer, sheep, fish*

The irregular plural realisations either have a different plural morpheme or a zero morpheme, or they result in completely new lexical items. In this study, however, I focus on the regular plural morpheme “s”. Interestingly, the “s”-plural is not such a regular and simple plural form as may be assumed at first.

Sometimes, a plural may not be marked at all because it is redundant in case of prior plural marking in the determiner, for instance. Rohdenburg (2004) has studied various low German and non-standard English dialects and found a tendency to delete the plural inflection on nouns after prior plural marking in a numeral, e.g. “three kid”. Interestingly, such dialects do not consistently use the unmarked plural; without a prior numeral the same nouns may carry a plural inflection in different syntactic contexts. Indeed, Dutch (not a low German dialect, of course) uses the plural “drie jaar” (literally *three year*) to refer to a definite period of time as in: “ze woont sinds drie jaar in Amsterdam” (she has lived in Amsterdam for three years) and “jaren” (years) to refer to an indefinite period of time e.g. “ze woont er al jaren” (she has been living there for years).

Apart from irregularities in using the plural inflection, the simple plural “s” morpheme in English may not only indicate a quantity of ‚more than one‘: It may also imply a generic sense, as in the sentence: *Cars are expensive.*

Apart from the binary number distinction between singular and plural, many other factors might be at work in determining the number marking of nouns. Such factors may be definiteness or countability. Obviously, the above example refers to an unspecified number of cars and involves a generic interpretation without pointing to a definite set of cars that may be visible. The distinction between mass and count nouns, however, seems to be a second dimension in addition to the singular / plural one and is therefore independent of the latter. Another factor influencing morphological number marking is semantic irregularity, in the sense that the plural is not just the singular in a larger quantity but it implies a specific semantic meaning. Such is the case in pluralised nouns as: *skies*, *heavens*, *looks*, *waters.* In this context, it might be interesting to note that the first verses of the Bible have been translated differently with respect to either singular or plural form of heaven (Pasierbsky, personal communication): *In the beginning God created the heaven and the earth* (King James version).

Heaven, in this case, has been translated from the Hebrew word "shamayim", which carries a dual inflection. In the course of history and depending on the target language, the translators provided different number markings, corrensponding to the inherent dual of the Hebrew word *heaven*. In other versions, "ha-shamayim" has been translated as "the heavens". Therefore, the marking of plurals may sometimes become a matter of semantic interpretation.

Not many languages mark the dual grammatically. Instead, the plural inflection often contains a conceptual dual, especially in the case of body parts. A language that has no dual form has to mark the conceptual dual with either a singular or a plural marker. Most European languages have a plural form and therefore distinguish everything that denotes one entity from everything that denotes more than one entity. Consequently, the conceptual dual is implied in the plural. The world´s number systems can be grouped according to a universal number hierarchy in the sense that every language that marks a plural form also has a singular form. Every language that marks a dual form also has a singular and a plural form (Croft, 1990). In languages that mark the dual grammatically, the dual number constitutes a form for anything that refers to a conceptual two-fold or a pair. In English, a language without a separate dual form, these pairs have to be marked differently by referring to a specific paired set or by using a quantifier: *a pair of hands*, *two ears*, *my eyes*, *two sequential days*. All of these examples would be marked with a dual marker in Hebrew, for instance.

The plural “s” morpheme in English is not restricted to a usage where it denotes anything more than one; it may also be used in cases where the referent clearly constitutes a singular entity. Examples include: *zero degrees, every means*, *a new scissors*, *I am friends with* (taken from Hirtle, 2000). For Hirtle, reference to either a singular or a plural concept is not the relevant factor for the functional distinction between the use of either the singular or the plural. The common element of meaning in the various uses of the “s” plural „appears to be an impression of discontinuity in space telling us that the quantity represented by the substantive is separate from other, adjacent quantities if one imagines a particular position on a scale” (Hirtle, 2000: p.61). Such is also the case in *zero degrees* which appropriately refers to the image of a scale that Hirtle used. In simpler words: “Pluralizing a noun takes its denotation from a set of individual objects to a set of collections of such objects” (Hurford, 1987: p. 42). Therefore, the common element of meaning in the use of the plural is directly related to such non-linguistic principles of boundedness and proximity as described in Gestalt theory (Wartheimer, 1938).

Similarly, Rohdenburg (2004) considers the absence and presence of plural marking in different contexts as examples of the different degrees in the universal hierarchy of individuation proposed by Sasse (1993). All these attempts to provide an explanation for the idiosyncratic uses of the plural inflection support a conceptual basis for plural usage. Nonetheless, individuation is also influenced by definiteness and therefore the presence of plurals in non-standard dialects depends to some extent on the determiners that are combined with the plural nouns (Rohdenburg, 2004). This shows that it is not merely the perceptual system that determines the choice of countability; language is also highly conventionalised, as Svensson (1998) points out in a dissertation dedicated to the question of defining countability. After an extensive overview of previous literature on this topic, he carries out a corpus analysis, providing many examples from various languages. He concludes that the number system is based neither on a purely arbitrary syntactic account nor on a completely iconic conceptual account. Rather, countability bears a lot of perceptual cognitive reality but cannot purely be explained by simplistic mapping from conception (Svensson, 1998). He concludes that the number category is largely based on the notion of embodiment. As such, it is closely tied to the way the embodied human being categorises the likewise embodied environment and makes some sense of it through sensomotoric perception.

4.1 The dual number

For an understanding of the semantics of the plural inflection, it is helpful to take a closer look at the functional distribution of the grammatical dual and the grammatical plural in a language like Hebrew. This is particularly important for an understanding of the second experiment reported in Chapter 8, where conceptual dual and conceptual plural forms are compared in processing. In terms of countability, the grammatical dual denotes everything that is paired; it consists of two singular entities in a bound context, and items that regularly appear together. A prototypical language containing all these number distinctions is Sanskrit (Tobin, 2000) where the dual constitutes anything paired and the plural anything „more than two“ (p. 96). However, not all languages containing a dual marker handle the semantic distinctions in the same way. Tobin states (2000: p.97) that the plural number in Hebrew may also contain the dual or anything paired, thus referring to anything “more than one“. Thus, dual and plural therefore exist in a relationship of inclusion. Likewise, the Slavic dual does not have the sharp semantic disinction that exists in Sanskrit. The common Slavic dual was a special case „since the common Slavic number category did not present a three-way distinction of singular vs. dual vs. plural“ (Janda, 2000: p.77) but rather used the dual number to distinguish between arbitrary pairedness and naturally occuring pairedness. Dual forms have also been impoverished in their morphology because the dual appeared without any other marking for gender, case or person (Janda, 2000).

Tobin (2000) questions the grammatical status of the dual number. He provides ample evidence for the lexical status of the dual number in Hebrew, thus competing with the lexicalized numeral „two“ in its compound form (e.g. two eyes {numeral *two* + pl noun} vs. eyes {dl noun}). In a way these two lexical entries with almost the same meaning (synonyms) compete for activation. In this case, the grammatical dual -like its gender counterpart (Sabourin, 2001)- would be on the verge of being lexical and therefore function as a productive member “as part of both the grammar and the lexicon simultaneously“ (Tobin, 2000: p.88). Similarly, Bantu languages like Swahili employ the number distinction not only to denote numerosity but also for categorising noun classes. As such, number is not purely grammatical but “closely tied to the problem of the meanings of the noun” (Contini-Morava, 2000: p.3). There is great complexity in the singular / plural distinctions for such languages and they do not show the conventional binary pattern as in our languages. Instead, number can be seen as a continuum of individuation. This is what makes the dual such an interesting category to study.

The key question leading this study with respect to the number category is how plural inflections are processed when different processing routes may be assumed depending on the conceptual nature of the plural. As stated previously, inflectional morphology is being associated with grammar and grammatical processing, and so are the plural inflections. However, number is not the same as gender, which has to be learned for each lexical entry in the learner´s lexicon. Number can usually be inferred directly from conceptual structure, unlike grammatical gender (Pienemann, 1998). In the case of the grammatical dual, an inflectional morpheme may add more to the word than just grammatical information. Sometimes, inflecting a word might lead to a subtle semantic difference. Unlike the Hebrew plural, the English plural incorporates both the dual meaning and the plural meaning but may be conceptualised in those two different ways. This is where grammar becomes also a lexical matter, directly inferred from concpetual structure. Hence, the plural can be defined as somewhere between grammar and lexicon. The discussion about the nature of the dual is important here, because I argue in my second experiment that the English plural may sometimes conceptually be understood as a dual. The dual´s ambivalent nature in Hebrew offers another argument for the borderline position of the English plural, as will also be confirmed experimentally in Chapter 7.

The dual, as Tobin (2000) argues for the Hebrew language, does just that. It semantically distinguishes the lexical item from its plural + numeral counterpart. Therefore, the composition is more than the sum of its parts and the inflection is more than just the application of a grammatical rule. The dual form, and in some cases the plural form, may be complete lexical entries with idiosyncratic meanings. Rather, the choice may depend on subtle semantic distinctions from its lexical counterpart (numeral two + plural compound). Since the use of the dual in Hebrew is highly idiosyncratic, restricted in number and limited to nominals only -also in adjective noun combination with a dual noun- as well as rather arbitrary in its distribution, it is hard to maintain a purely grammaticalized view of the dual number. The dual is also neutral in gender and it is not part of a tight paradigmatic system since the singular and plural can in certain cases co-occur with the dual in the same lexical item and they do not exclude each other. Not all items appearing as pairs or paired sets in our environment are marked with the dual; the distribution seems arbitrary in the Hebrew language. Semantically, the dual number is not always straightforward; it can even function as a plural designating more than two referents. Therefore, the semantic domain of countability as the basis for number decisions is questioned. Tobin (2000) prefers to use the concept of *semantic integrality* to define the often subtle differences in the use of dual and plural number in Hebrew. He provides examples of nouns with dual inflection in one case and numeral (*two*) plus plural inflection in the other case and describes the subtle semantic differences between the two. Similar to the Slavic languages (Janda, 2000) and similar to the the discussion about continuity versus discontinuity as the crucial factor in distinguishing dual and plural (Hirtle, 2000), Tobin claims that the dual morphology is marked with respect to semantic integrality. It only allows for an integral interpretation ("a dual unit") while the unmarked form prefers a discrete ("two entities"), non-integral interpretation similar to the difference illustrated here: a+b (discrete entities) versus [ab] (continuous integral, thus dual). Being a marked form, it is considered to be cognitively more complex than its unmarked counterpart.

The strict distinction between lexicon and grammar cannot be sustained. Here, it may be interesting to note that von Humboldt (1836), in his essay on the nature of language, already stated that the distinction between grammar and vocabulary merely serves practical purposes in teaching a language. For him, the form and substance of a language cannot easily be separated since form influences and therefore affects the substance. He argues that language form is in some ways connected with the lexicon itself, it reflects the intrinsic nature, and therefore also the national spirit, of a language. The last comment is, of course, to be read historically as part of the justification for the definition of nation states through distinct language. In the light of recent research on the nature of grammar and lexicon with modern research techniques this insight certainly seems very advanced for its time, even though it does not exactly follow the same argumentation as today. More recently, however, Schönefeld (2001) has addressed the issue of interdependency between grammar and lexicon and found evidence from speech production data as well as from experimental research for the inseparability between lexicon and grammar in language processing. How conceptual categories such as dual and plural influence linguistic processing will be explored in the second experiment of this study. Even though the experiment uses only pluralised English nouns as linguistic stimuli, the underlying conceptual nature will be seen to have an effect on processing.

4.2 Numerals

Numerosity is expressed not only by pluralising nouns and verbs; it can also be expressed in numerals. Numerals are an integral part of the stimulus items in the second experiment of this study; I discuss them briefly in this section in order to explain their potential impact on processing the stimulus. Verbal numerals ("two", "three") are acquired at a very early age. Number words are recognized as referring to precise quantities from very early on (Freeman, 2000). For monolinguals, it is therefore to be assumed that such numerals are stored in the mental lexicon as exactly referring to a corresponding conceptual quantity. This numerosity may be very concrete in case of the lowest numerals (e.g. *two* or *three*) and have a more abstract quality in the higher regions (e.g. *five million two hundred and twenty-six thousand*). In the latter case, the speaker does not have the exact number of entities in mind; the number has rather become an abstract quantity (Hurford, 1987). The fact that numerosity is concrete in the lowest numerals is also proven in processing experiments and is generally referred to as "subitizing" because low numerosities (up to four) can be directly named without overt or covert counting. When the quantity increases, the speed or accuracy decreases. This number-size effect and other number related qualities can be directly transferred to the symbolic representation of quantity in numerals or digits (Dehaene, 2001). Numerals and other symbolic representations of quantity therefore directly refer to their underlying concept and there seems to be "an internal access to quantity" (Dehaene, 2001: p.8). Processing numerals for such small numbers might therfore not very likely show in reaction time since direct conceptual mapping is involved. Therefore, reaction times between small numerals and plural nouns should merely reflect the processing time of grammatical agreement and not of quantity computation.

Numerical symbols, whether the Arabic number symbols or the word symbols in Roman letters, refer to one and the same underlying nonverbal numerical semantics of the symbol (Whalen et al, 1999). In terms of nonsymbolic numerical perception, number perception is not an exclusive human skill; animals are also capable of scaling numerosity.

The nonverbal representation of numericals is generally believed to be a scalar variability (Dehaene, 2001; Whalen et al, 1999). Whalen et al have demonstrated through empirical data that the perception of numerosities is independent from overt or covert verbal counting or the perception of duration. Verbal numerals and their conceptually inherent numerosity refer to a scalar system, just like Arabic number symbols, and their processing is not influenced by their symbolic character. This is also important for the outcome of the reaction times in the second experiment: small numerals have the same processing quality as their number symbols.

4.3 Processing linguistic numerosity

The processing of numeric information depends very much on the function that numerosity expresses linguistically. First of all, it seems irrelevant whether numerosity is expressed by a number word or an Arabic numeral, as reported above. In a linguistic environment, however, certain factors do impede with this direct conceptual mapping of numerosity. Baayen et al (2003) have found a very stable number effect for singular versus plural nouns, where singulars are consistently recognised faster than plurals in on-line experiments. However, these effects could not be replicated in the same experiment with verbs. One explanation might be that the nouns directly refer to underlying conceptual scalar numerosity, as explained above. Verbal plurals reflect grammatical number and do not have independent access representations (Baayen et al, 2003) but depend on linguistic context and agreement procedures. Even pluralised nouns are so straightforward due to the subtle semantic changes that take place through pluralisation. The influence of extra-linguistic factors in linguistic processing will be demonstrated in the two experiments of this study where processing of pluralised nouns is being compared in different conditions. In the second experiment, pluralised nouns co-occur with a numeral determiner. The subjects are asked to take a grammatical judgement decision. Subjects will be presented stimuli such as:

- *many guest* vs *many guests*
- *one glasses* vs *one glass*
- *two foot* vs *two feet*

The two conditions crucial for this study are the nouns that refer to conceptual duals e.g. *feet* in comparison to nouns that refer to conceptual plurals (usually more than two) e.g. *drops*. In the case of the conceptual duals, it is hypothesised that the nouns are processed faster in combination with a dual numeral (*two*) than with a plural numeral (e.g. *three*). It will be interesting to see whether there are differences between e.g. *two hands* and *many hands*. The reaction time could differ between the two determiners for reasons discussed above: the numeral „two“ being conceptually represented as a dual, the second not referring to any specific kind of number. Second, reaction times could differ because the plural dominant (the plural occuring more often than the singular) lexical representation of *hand* might be stored in the lexicon as a conceptual representation of a dual, in which case the agreement with „two“ should be more straightforward than with other numerals or determiners. Finally, the experiment will prove that conceptual factors play a major part in processing grammatical plural inflections, as already hypothesised by Baayen et al (2003) and observed by Rohdenburg (2004) and Svensson (1998).

5 *Morphological processing*

Now that the different semantic functions of the plural and the complexity of the number system have been scrutinised, it is necessary to take a closer look at the regular plural inflection and its processing nature as reported in the literature. The fundamental debate between a purely grammatical account versus a conceptually induced account of pluralisation will be reflected in the following sections. The experimental evidence summarised in this chapter will eventually support the idea of a gradual distinction between grammar and lexicon, reflected in the processing nature of plural inflections.

The assumption that morphosyntactic encoding depends on different sets of processing routines is the basic tenet of this project. As will be shown in this chapter, even within the realm of one type of inflectional morpheme -here: the plural "s" morpheme- various processing paths are probable. The following sub-sections will raise the question of how to define morphological representation in the mental lexicon, focussing on plural inflections. As will be seen by the end of this chapter, there is hardly a unitary processing account for inflectional morphemes, which may question the very existence of the inflectional category in processing terms. It is therefore necessary to review psycholinguistic evidence that may tell us something about the mental representation of such inflectional knowledge in a more specialised way than already introduced in the first chapters. As mentioned at the very beginning, the fact that speakers have acquired inflections does not tell us anything about the nature of the processing involved. Psycholinguistic experiments reflect the nature of inflectional processing by means of reaction time. The time course of linguistic processing in experimental settings may reveal the order of possible subroutines in language processing. But before we can define the processing procedures for inflectional morphology, we must clarify why inflectional morphology is considered a grammatical category in its own right.

5.1 Inflection versus derivation

Whether inflectional morphology is an inherent part of the lexicon or whether inflections - as opposed to derivations- are processed post-lexically is a hotly debated issue (e.g. Perlmutter, 1988; Badecker & Caramazza, 1989; Schriefers, Friederici, & Graetz, 1992, Tuggy, 1997, Baayen, Dijkstra & Schreuder, 1997). It is not so easy to say whether inflections are lexical or grammatical. Although this question may seem clear from what we know from grammar textbooks, this question may have another dimension in terms of on-line processing. Some psycholinguistic evidence indicates a separation of both morphological functions in native speakers (Hancin-Bhatt & Nagy, 1994), and therefore for split morphological processing. Inflections are of a purely concatenative nature in the English language and appended at the end of a word. But, as research in American Sign Language (ASL) shows, this formal property is not the reason for their idiosyncratic processing route. In ASL, inflections are non-concatenative and are encoded parallel to the lexical content information (Emmorey, 1995). Nevertheless, ASL processing experiments suggest a specific processing route for inflections (Emmorey, 1995) as opposed to derivations. Despite the different modes of language input and the different nature of inflections, a vast amount of research with healthy and aphasic patients supports the idea of an independent morphological structure for inflections (Badecker & Caramazza, 1989; Schriefers, Friederici, & Graetz, 1992). This distinction is also reflected in the processing differences between regularly inflected and irregular forms (Sonnenstuhl et al, 1999; Lück et al, 2001; Münte et al, 1999). The way lexical and grammatical encoding is thought to take place in speech production was summarised and modelled by Levelt (1989) and has already been explained in Chapter 2.

I have already shown that the class of inflectional morphology is not as clearly definable. For the case of English plurals, forms like *scissors* and *trousers* do not have a singular form (although they may grammatically agree with the singular: *Where´s the scissors?*) and are complete lexical entries even though they are inflected as plurals. The semantic nature of the inflected lexical item seems to play a role in processing. Booij (1993) ascribed different processing behaviours to inherent inflections (of a more semantic nature) and contextual inflections (due to formal properties). This distinction has also been observed by Tobin (2000) and Hirtle (2000) and Hay (2001) who conclude that the plural has an ambiguous nature somewhere between lexicon and grammar. Pienemann (1998) distinguishes between different kinds of plural processing depending on an inherent lexicalised plural or a syntactic plural agreement. Even within the inflectional morphological paradigm there is evidence for different processing procedures. A strict split morphology model with separate processing routes for inflection and derivation seems problematic.

Other evidence for a more gradual distinction between grammar and lexicon comes from research on frequency effects in the processing of plural forms. Frequency has shown to be a very important factor for the processing nature (Sereno & Jongman, 1997; Davies et al, 2001; Baayen et al, 1997). Not only frequency but also plural dominance (i.e. plurals occurring more often than their singular counterparts) seem relevant (Levelt et al, 1999; Baayen et al, 1997). The inflectional plural-ending on nouns cannot be explained by just one processing path. The question of a more unitary processing account depending partly on lexical input (the cause of frequency effects) and causing processing gradience becomes pressing.

Little is known about how the above factors influence the processing of inflections in second language learners. If native speakers clearly show frequency and dominance effects, as well as semantic plural effects, as they process plural words, these effects should have a point of emergence in language learners and gradually show up more strongly. This study sets out to determine the exact nature of these influencing factors and determine their point of emergence in language development.

5.1.1 The place of inflections in language representation

In order to define the class of inflectional morphemes and to understand why traditionally it has been seen as a clear-cut class as opposed to derivational morphemes, it is necessary to know what is normally considered an inflectional morpheme.

Speech errors and slips of the tongue in everyday language suggest separate representation of stem and affix in word production. Errors like morphological stranding: „vor Pauschal**ver**hüt**ung** sollten wir uns urteil**en**" (personally heard on the radio in February 2002) clearly show that speech errors follow certain morphological rules that indicate separate encoding of stem and affix. Morphological stranding means that the morphemes are stranded in their originally intended position; only the lexical stems are exchanged. In the cited example, the derivational morphemes ´ver´, ´ung´ and ´en´ have not been moved from the original position, but the speaker exchanged the two lexical word stems: 'hüt´ and 'urteil'. Such unintended errors give us some indication of the nature of word formation processes in speech production. Levelt (1989) gives an example of inflectional stranding: "a hole full of floor**s"**. Here, the inflectional ‚s' is stranded at its original position whereas the corresponding nouns are exchanged. This example shows that the inflectional ´s´ is a separate category and not inherent to the word stem; therefore it is not stored as part of the lexical item. In morphological shifts like „they needed to be mad**ed**" (taken from Levelt, 1989), the inflectional ´ed´ suffix is repeated in a position where it was not intended. This also demonstrates that the inflections are a separate morphological category not adjoined to the lexical stem.

Shattuck-Hufnagel (1983) attempted to answer the question why sublexical elements (elements smaller than the word level) should participate in exchange errors at all. She assumed that expected speech errors would concern the serial ordering of words (syntax) and not intra-word structures, since we expect words and not whole sentence strings to be retrieved as complete units from the lexicon. However, her study of a speech error corpus shows that most spontaneous speech errors are in fact of a sublexical nature. Speech production should therefore contain a mechanism of a serial ordering process for sublexical elements. Shattuck-Hufnagel mainly found that phonemes were the smallest sublexical units to be exchanged, though larger units, such as morphemes, can be subject to exchange errors. Thus, morphemes can be regarded a category separate from the lexical category.

As mentioned at the beginning of section 5.1, at least for the English language, all inflections are realised as suffixes and therefore follow the word stem. For other languages, inflections may also be realised as pre- or infixes, therefore preceding the word stem or inserted in the word stem. In typological comparison, inflections seem to be an optional medium to express grammatical relations. As Klein (2000: p.3) says: „In the design of human language, inflectional morphology is a common but by no means indispensible part". Some languages, like Chinese, have only a marginal inflectional system, if any. If the human language faculty does not require inflections and if they are not a necessary cognitive component of language, it is legitimate to ask whether their processing displays special properties that are not relevant for derivational morphology. Inflections in the English language, for instance, cannot take any further suffixes and usually appear further away from the word core than other elements such as derivations. For instance, *swims-ness* is impossible. Usually, inflected forms cannot be the first word of a compound although in some languages such combinations do occur. In English, it is impossible to create something like **swims-suit.* All other morphemes can take other suffixes, though.

Various factors in spoken language support the notion that stem and inflectional affix are represented at different places. This is also illustrated in Janssen's (1999) extension of the Levelt/Roelofs speech production model that assumes a split morphology processing procedure. Janssen adds an independent morphological encoder between grammatical and phonological encoding. The encoder distinguishes between stem encoding and inflectional encoding during lexical processing. This model assumes that word stems in speech production are represented in the mental lexicon and need to be accessed from the lexicon whereas inflectional morphemes have to be encoded post-lexically after lexical access. The morphological information of the lemma is already in the lexicon, however, and added in the form of diacritic tags that form the morphological spell-out of a lemma. This means that the possible morphological realisations of the word stem as well as its syntactic functions are already spelled out in the lexicon in the form of unspecified frames added to the word stem. The frames are filled during speech production by means of separate procedures or subroutines. This “slot and filler mechanism” of unspecified inflectional frames in speech production has been proven in a set of psycholinguistic experiments (Janssen et al, 2002). As Janssen et al (2002) state, such inflectional frames may also explain slips of the tongue and speech errors. In the case of speech errors, the inflectional frame is spelled out but erroneously filled with an inappropriate filler. In case of regular plural inflection, the inflectional slot is filled by using a rule-based grammatical subroutine. Such inflectional frames differ from language to language. English nouns, for instance, can carry stem and number information. German nouns, additionally, carry gender and case information that needs to be realised morphologically. The lemma with its diacritic features is then transported to the morphological encoder where the morphemes are generated.

5.1.2 Split morphology hypothesis

The above view of morphological encoding clearly does not consider inflected words to be part of the lexicon. This idea has often been adapted in linguistic research and is usually known as the split morphology hypothesis, put forward by Perlmutter (1988). The split morphology hypothesis, as Mc Queen and Cutler (1998) defined it, assumes that regular, productive inflection is extralexical. Exceptions to the split morphology hypothesis are inflected forms that are stored in the lexicon as separate lexical units, e.g. *clothes.* „Clothes“ is a monolithic form that does not have the same meaning as its uninflected form *cloth.* The fact that lexical irregular plural forms may be inflected in some non-standard dialects or even in child language acquisition shows that the irregular form is a complete lexical item. In the split morphology hypothesis, irregular forms are lexical just like derivational ones. Children frequently overgeneralise the plural rule and produce words like *childrens*. Such redundant plural marking can also occur in non-standard varieties. For example in a northern German dialect „Kinders“ or „Mitarbeiters“ are possible plural forms; on a TV program in January 2003, a farmer used these forms when he was interviewed. This seems to suggest that the default plural encoding is based on regular syntactic rule application and all other forms are exceptions. This view has also been proposed by Anderson (1992) who says “We thus take the formal constituents of complex words to be not listed morphemes, but rather operations on the form of words”. He continues by demonstrating this idea with the example “dogs”. In itself it results not from a formation process of two morphemes listed in the lexicon but is rather a form “dog” that has to undergo a rule which results in a morphological and phonological change and represents the property [+plural] (Anderson, 1992). Such a dichotomy in morphological processing is suggested in the split morphology hypothesis.

So far, there has been an acknowledged distinction between stem and inflectional affix. But which morphological forms are inflectional? Traditionally, grammar books define inflectional morphemes as not changing the word class of the wordstem. If inflections are syntactic and derivations have a lexical quality, this would be an argument in favour of the separate processing organization of such categories as hypothesised in the split morphology account. The nature of inflections cannot be defined solely by reference to productivity since some inflections, like the plural form of "children", are not productive despite their agreed inflectional nature; on the other hand, derivations such as "-ing" may be highly productive, as well (Anderson, 1992). Anderson points out that productivity and generalisability cannot solely explain inflections (1992). Likewise, in defining the inflectional category, it is not enough to just list the features that are traditionally considered inflectional versus those that are traditionally considered as derivational. As Anderson (1992) points out, such differentiations may not be universal. Some languages may treat a feature in a more inflectional way, whereas other languages treat it derivationally. Such is the case with diminutives. It is clear, though, that the properties considered inflectional have one thing in common: the information exchange between the word and its syntactic environment, thus agreement features, phrasal properties, inherent properties (like gender) and configurational properties are typically inflectional (Anderson, 1992). But it is nevertheless difficult to exactly pinpoint the distinction between the two categories and a split morphological account for morphological processing is hard to sustain.

5.1.3 Inflections in learner language

It is only plausible to conclude that inflection follows language-specific rules and has to be acquired by the second language learner. If inflection is driven by syntax and is, thus, language specific, how can it be accounted for in acquisition? In a psycholinguistic framework, a traditional distinction between inflection and derivation should also be reflected in processing; otherwise such a classification seems not to be justified in processing terms. The suggested processing difference between inflectional and derivational morphemes in language processing data has been demonstrated by various psycholinguistic experiments like that of Hancin-Batt and Nagy (1994). They demonstrated that second language learners of English have more difficulty translating suffixed words than simple word stems in the case of non-cognate translations (into L2 English). However, the difficulty was much more apparent in stems with a derviational suffix than in stems with an inflectional suffix.

First language acquisition studies of English native speakers have shown that children master inflectional morphology very early, often before they finish acquiring derivational suffixes (Berko, 1958). Such observations support a processing difference between derivational and inflectional morphology. However, even in first language acquisition, inflections are not self-evident from the very beginning. The fact that a child has mastered the inflectional form of a plural, for instance, does not mean that it understands the fact that the inflectional ending denotes a plural (Frijn & de Haan, 1994). The understanding of the inflectional ending as denoting the semantic plural occurs around age four and is one of the first inflections to be learned cross-linguistically (Frijn & de Haan, 1994; Goorhuis & Schaerlaekens, 1994) The acquisition process shows in overgeneralizations or backformations, when children erroneously analyse word forms as decomposable into stem and plural suffix. For example, *"een vark" (child aged 3.1) is a non-existing word and results from the erroneous assumption that the word „varken" (=pig) is a plural form because the „-en" ending is a regular plural ending in Dutch (Frijn & de Haan, 1994). As Tomasello (2000: p. 215) puts it:

> ...data-intensive studies from a number of different languages together show a very clear pattern. First, young children's earliest linguistic productions revolve around concrete items and structures; there is virtually no evidence of abstract syntactic categories and schemas.

Children have to gradually learn to analyse and decompose inflected forms just as learners of a second language do. An investigation of the longitudinal data of a native Arabic speaker learning Dutch, as documented in the ESF (European Science Foundation) corpus of natural second language acquisition (McWhinney, 2000), revealed that even after a long time living in the target language country, learners can have severe difficulties with the simple, productive and very frequent plural inflections; very often they just leave them out. This observation is also confirmed by van de Craats (2002) who reports on two L2 learners of Dutch from the same ESF corpus. Both the Turkish and the Moroccan learner either completely leave out plural affixes or erroneously use pluralised words as chunks when they refer to the singular. When a learner occasionally uses the right plural inflection, he may drop it again in the next instance. So even though inflections have been assumed to be easier to acquire than derivations in a second language (Hancin-Bhatt & Nagy, 1994), L2 learners -at least in natural acquisition- frequently just drop inflections. But even in L1 acquisition, the correct production of plural forms does not necessarily imply that the child has mastered the plural system. The child may just as well consider the plural form to be a lexical unit and therefore a whole-form representation in the lexicon (Frijn & de Haan, 1994).

5.1.4 Different processing accounts

The question of whether the morphological classes can really be defined contrastively remains debatable. Another possible processing account might be a gradual distinction in morphological representation (Bybee, Kurytowicz, Plank ,1995). This assumption is also advocated by Booij (1993) who claims that there is no such thing as ‚split morphology‘. That is, he says that morphology cannot just be divided into an inflectional and a derivational class; he argues for a distinction between inherent and contextual morphology and, therefore, for an alternative way of categorizing morphology. Inherent morphology, he argues, is the morphology that is semantically attached to the lemma, therefore intrinsic to the word and stored as such in the lexicon. Contextual morphology, being a purely syntactic construct, is not semantically relevant and its functions are peripheral. Contextual morphology, like agreement inflections, is not lexicon-intrinsic. Booij´s distinction also presupposes a dual processing route. It does not, however, conform with the conventional classification into derivational and inflectional morphology. Rather, Booij‘s classification is of a semantic nature and not functionally induced. This is in line with the Processability theory of language acquisition developed by Pienemann (1998), who found evidence for such processing differences in learner language.

Booij (1993) argues that the plural is an inherent inflection, thus semantically inherent and not triggered by syntactic context. Therefore, the inherent inflection is qualitatively different from the contextual inflection and has a more derivational character. Both agreement morphology and lexical morphology, however, may be inflectional, as for instance in the case of the plural inflection. The distinction in this case is more of semantic type. Booij provides several instances of plural use that seem to support the notion of an inherent semantic nature. He argues that plural agreement inflection is contextual and in a way redundant because plurality in a noun phrase is already marked in the determiner (e.g. *two children*). Similarly, Tobin (2000), Contini-Morava (2000) and Hirtle (2000) argue that the plural is ambiguous, being at the intersection between syntax and the lexicon. The literature reviewed above seems to support a tendency towards a more semantically induced distinction between different morpheme groups, and therefore also different processing routes. Even though this processing account also presupposes a dichotomy of processing routes, it is based on a fundamentally different explanation. Such a functional difference of processing routes has been accounted for in acquisition data (Pienemann, 1998) and may also explain the processing differences between the traditionally described morphological classes.

Hypothesis: There is evidence for different processing procedures within the realm of plural inflection. These different processing procedures may be a result of inherent versus contextual morphology or of lexicon-intrinsic (e.g. *clothes*) and inflected plural forms or due to other extra-lexical factors.

5.2 Processing differences due to formal morphological properties

A dual processing route based on inherent and contextual morphology may give rise to serious doubts about the distinct and clear cut nature of inflectional morphology. If the processing of inflectional morphology differs from that of derivational morphology this may not come as a surprise. Very often, inflectional morphology is realised by suffixation, at least in the English language. So far, however, most research in this area has been conducted on language processing in English or the closely related Germanic languages. Processing differences due to the formal properties of inflectional morphology must be considered a possibility. Suffixes are added to the word stem; they are therefore serial in their appearance in linguistic encoding and reception. If the word stem and its suffix need to be processed in comprehension, access times may differ because of the increased length of the word or because of prior access to the word stem proper before detection of the suffix. Such explanations need to be considered when interpreting the above findings of dual processing accounts.

5.2.1 Language mode

At the beginning of section 5.1, I referred to experimental evidence for dual route processing from American Sign Language (ASL) (Emmorey, 1995). Unlike spoken language, sign language usually encodes morphological information non-concatenatively, i.e. in a parallel way and simultaneous to the word stem. This typical characteristic of sign language found cross-linguistically is already constrained by the simple fact that signing is much more time-consuming than speaking since an isolated sign lasts twice as long as the an isolated spoken word (Emmorey, 1995). Therefore, it is cost-effective to encode as much information as possible in a simultaneous manner.

Research on ASL is important because it reveals a lot about the neurolinguistic nature of morphology despite the different linguistic mode. Since ASL encodes inflectional information in a non-serial way, processing of inflections in ASL does not depend on serial processing. If the effects of inflectional processing in spoken English occur only because of the seriality of the inflection (the English inflection being a suffix), the nature of the processing may just be due to formal properties. If research results on the inflectional nature in ASL differ from those on spoken language, it is reasonable to assume a modality effect. The processing route of inflections might therefore depend on the mode of language input and not on the inflectional nature. Because inflections in spoken language are mostly of a concatenative nature, it may be natural to assume a temporally delayed processing for inflections with prior recognition of the base form (Emmorey, 1995). The dual-route model addressed in the previous section might then be due only to the formal morphological properties and not to the morphological nature. As opposed to non-concatenative encoding, a continuous encoding in spoken language is associated with less computational complexity; given the high speed of spoken language, this is a clear evolutionary advantage.

In the light of previous research on inflections (Emmorey, 1987) in auditory versus visually presented language, the assumption of formal processing differences would be perfectly justified since the processing route - at least for English suffixes - depends on the input mode. Inflected words may be processed in a parallel manner when presented visually but serially when presented auditorily. Interestingly, however, the ASL studies show that despite simultaneous encoding of stem and inflections, perception is not simultaneous, at least for verbal inflections. Even though they are presented in a parallel way, stem and inflection need separate perceptional attention, to be measured in reaction time. All this suggests an independent morphological structure in both spoken and signed language and therefore supports a neurolinguistic justification of a dual-route model for inflections. This view, of course, presupposes a division between inflectional and derivational morphology as it was presented in the previous sections.

5.2.2 Typological differences

Other formal properties that might influence morphological processing are the cross-linguistic differences in inflectional realisation. As Klein (2000) points out, most inflections in English are mere additions to an otherwise complete word stem, as discussed in the previous section. In Italian, the removal of an inflectional element may result in an incomplete word form. For instance, if the inflection -*iamo* is removed from "bev-*iamo*", the remaining root is a bound form. It is therefore difficult to imagine that all inflectional morphology should be processed in the same way, namely via stem retrieval and affix computation. For inflected words in Italian, full form representation seems likely; it was demonstrated in a study by Say and Clahsen (1999). Interestingly, not all Italian inflected verb forms seem to be stored as wholes; the most frequent and therefore also most productive forms -those based on the 1st conjugation paradigm, ending in -*are,* seem to be rule-based. This first conjunction paradigm is by far the most frequent for Italian verbs. The others, infinitives ending in -*ere* and -*ire*, are rather infrequent and also not productive, a finding also supported by learners´ overgeneralization errors and L1 acquisition data (Berretta, 1995). This may indicate that frequency and productivity play decisive roles in the way morphemes are processed despite the morphologically ambiguous and seemingly lexical form of the inflected Italian verb.

The same conclusion can be drawn from a study that investigated the nature of Finnish inflections (Bertram et al., 2000). The question leading this study was whether homonymous inflection affects the processing procedure. In morphologically rather poor languages affixal homonymy has been found to influence processing. Inflected words with homonymous suffixes are stored as full form representations in order to avoid computational confusion in the parser because homonyms present multiple possibilites. In the Finnish language such a full store processing behaviour was considered unnatural because of the morphologcally rich input and the often multiple inflectional endings of

Finnish words. Besides, the relationship between inflectional form and its function is more straightforward in Finnish than in a morphologically rather poor language like German (Pienemann, 1998). As Pienemann points out, these different degrees of complexitiy in the relationship between form and function across languages, lead to substantial learning problems in second language development. Inflectional endings in Finnish, even homonymous ones, are assumed to be processed computationally because every Finnish stem can take a myriad of possible inflectional ending combinations. Because of this, each individual surface word form is relatively infrequent. The likelihood of very infrequent forms being stored as whole-form representations was considered very small. However, Bertram et al. (2000) found that even Finnish inflectional homonyms behave like those in morphologically poor languages like English or Dutch. The lexical decision tasks that controlled for base and surface frequency very clearly demonstrated that the Finnish language processor would also encounter problems in discriminating between the possible alternatives for homonymous affixes. Therefore, a whole-form lexical representation also seems to be more efficient for Finnish native speakers. However, as in Klein‘s (2000) observation above, this only applies to relatively infrequent and unproductive variants of Finnish inflected forms. As soon as a form becomes frequent and productive, computational processing is preferred. In such cases, computation may be additional and parallel to simple word retrieval and may therefore facilitate retrieval, as shown by Bertram et al (2000).

In the following sections, the existence of inflectional morphology as a separate class will therefore be taken for granted. This justifies research on inflectional morphology as a specific morphological class that is not merely different due to its formal properties (serial and suffixed) in the English language.

5.3 Regular and irregular inflection

As already seen in previous sections, there is ample empirical evidence for a dual route processing account. It was also shown that the underlying theoretical explanations may vary. Dual-route processing can be a consequence of two different ways of approaching morphological classification. In this section the main factor shown to induce processing difference will be the distinction between regular and irregular inflectional morphology. This distinction has been supported by several researchers working with psycholinguistic on-line experimental techniques (Lück, Hahne & Clahsen, 2001; Sonnenstuhl, Eisenbeiss & Clahsen, 1999; Münte et al, 1999). The most important work will be reviewed in the following section.

Using a cross-modal morphological priming task on noun plurals in German native speakers, Sonnenstuhl, Eisenbeiss and Clahsen (1999) demonstrated that a substantial difference between the reaction times for irregular and regular inflections. Since previous priming experiments on inflectional morphemes yielded inconsistent results for any kind of processing model, Sonnenstuhl et al repeated an existing experiment with an improved methodological design while assuming that regular morphological forms are processed by means of computations of decomposed representations of stem and affix whereas irregular morphological forms are full-form representations in the lexicon. Similar conclusions have been drawn from studies with aphasics where a dual-route processing model was favoured to explain processing differences. One drawback of previous priming experiments was the simple fact that they were often based only on the English language which has an obvious disadvantage when it comes to irregular inflectional forms. Irregular plurals in English, for instance, are very rare and supposedly represented as full-form lexical units because the word stem completely changes. In German, however, irregular plurals constitute the majority of all plurals and they are usually realised by suffixation[6]. In sum, there is a higher probability of ascribing any potential priming differences in German experiments to a processing difference between regular and irregular forms. For English plurals, one could argue that processing differences might be due to statistical distribution or due to a matter of form discrepancies with respect to morphological decomposability (the irregular usually not having any inflectional suffixes).

In their experiment, Sonnenstuhl et al (1999) presented regular and irregular word stems on a screen. They were acoustically primed by their respective plural inflected forms in order to avoid any modality-specific access priming effects. All conditions were compared to the base line condition where prime and target are identical. The reaction times show a full priming effect for the regular German -*s* plural, but not for the irregular German -*er* plural. One could argue that this effect is due to a frequently occurring vowel or syllable

6. it is being disputed whether the German ‚s‘ plural is to be considered the regular form and thus the less frequent variant or not.

structure change in the word stem when inflected with an irregular -er plural suffix which makes the noun *Haus* change into *Häuser* when pluralised. However, these factors were accounted for in the control priming task with diminutives where vowel and syllable structure changes are the only confounding factors. As these factors yielded no priming effects they were avoided in the plural task (cf. Sonnenstuhl et al. , 1999). The findings in this experiment clearly favour a dual-route model of morphological processing.

5.3.1 Evidence from neuroimaging research

For German plurals, similar results with different processing mechanisms for regular and irregular plurals have been found by Lück, Hahne and Clahsen (2001) in a comprehension task. Subjects were presented sentences with either regular (e.g. *Bonbons*) or irregular (e.g. *Kapuzen*) plurals or incorrect (e.g. **Kapuzes*), or irregularized regulars (e.g. **Bonbonen*). The two latter types represent violations of the German plural rules. Subjects were exposed to sentences with the pluralised stimulus words and their event-related brain potentials (ERPs) were recorded at the same time by electrodes on the skull. Incorrectly irregularised nouns (e.g. **Bonbonen*) evoked a LAN/P600 pattern, which is similar to the positivity observable while processing syntactic violations. The incorrectly regularised plurals (e.g. **Kapuzes*), however, evoked general negativity, quite the opposite of the incorrect irregulars. This could mean that the violation of a regular plural rule (e.g. **Bonbonen*) is experienced as similar to a syntactic violation, whereas the violation of an irregular plural rule (e.g. * *Kapuzes*) is conceived as similar to a semantic violation. This would support the idea of the regular plural ‚s‘ as a rule-based, post-lexical inflection. The violation of an irregular plural, perceived as similar to a semantic violation, would support the idea of failed access to the lexical item. The results were interpreted as illustrating different neural activity for German regular and irregular plurals: the irregular plurals as the lexical ones with full-form representations in the lexicon and the regular 's´ plural representing a rule-based inflection.

Whether the German plural rules can simply be taken as examples of a general dual processing of plurals is questionable. After all, English has only a minimal distinction between regular and irregular for inflectional plurals. Very often the irregular plurals are single lexical items and are therefore assumed to be lexical entries, anyway. What is interesting in this study is the very fact that a dual mechanism seems to be at work, one that can be measured in the quality of brain activity. A similar observation has been made by Ullman et al (2000) who recorded the ERP‘s of subjects shown sentences with and without violations of syntactic phrase structure and regular and irregular morphology. As in Lück et al (2001), violations of regular morphology evoked a different electrophysiological pattern than violations of irregular morphology. The study by Ullman et al. focussed on regular and irregular past-tense morphology rather than plurals but the pattern seems to repeat itself. Although the experiments described were comprehension studies, a similar distinction between processing mechanisms can be assumed in language production.

Such clear arguments favouring a dual route at least for the processing distinction between regular versus irregular plurals seem to provide no reason for doubt. The above findings were also supported by similar experiments focusing on the difference between regular versus irregular inflections, for instance in the English past tense. Münte et al (1999) found differences between regular and irregular past tense forms in a visually presented repetition priming task. They recorded the ERPs of their subjects´ performances in a primed lexical decision task. The underlying idea was that decomposition of regular past tense forms into stem and affix should show in the brain potentials, whereas irregulars are processed as whole-word entries with correspondingly different brain activity. The priming task presupposed that regularly inflected past tense forms would induce a priming effect. That is, reduced response latencies when presented prior to their corresponding stem forms whereas irregulars would not cause such an effect, as in the priming task of Sonnenstuhl et al (1999). In terms of brain potentials, this priming effect shows in a reduced N400 paired with a long-lasting late positive-trending shift between 500 and 800 ms after presentation of the target word. Like Sonnenstuhl et al, Münte et al (1999) also recognised the potential drawback of this study because of differences in the degree of formal similarity between prime and target for the two conditions (regular past versus irregular past tense). However, those problems were considered in the design of the study. The time lag between prime and target was unusually high in order to prevent immediate formal priming. They also varied the letter case for prime and target. They chose control items that shared the same initial syllables and would therefore be liable to induce formal priming. As hypothesised, the above-described priming effect holds only for regularly inflected primes. Also, there seems to be a minimal priming effect for regular nonce words, which suggests a kind of formal priming effect due to similar froms in the uninflected and the inflected form. The irregular priming condition shows a similar weak late priming effect. The irregular nonce word and the phonological control conditions did not yield any priming effect. Münte et al also measured the ERP‘s for unprimed past tense forms and compared them to their stem forms as well as the unprimed forms of the corresponding controls; they found differences in processing only in irregular verbs. This experimental sequence supports a dual-processing route for regularly and irregularly inflected verbs.

Also very interesting is the finding that the reaction times of the lexical decision task did not yield any priming effects for either condition. The priming effects were only measurable in the electrophysiological differences in brain activity and they could be reported only because of this particular experimental design with electrodes. The lack of any reaction time effect was explained by design aspects of the study (cf. Münte et al, 1999). More important, formally related primes and targets in the control condition yielded no priming effect at all. Sheer formal similarity, partly ruled out because of the specific design, can therefore not be the cause of the priming effect observed in the regular verbs. The difference between regular and irregular processing is not due to the degree of formal similarity. Although such electrophysiological processing differences in a priming task so far have only been found for verb inflections, it is very likely that they appear for noun inflections, as well. This is especially the case since the priming effect proves to be a morphologically intrinsic effect and it is not caused by mere formal properties of prime and target. In this case, morphologically intrinsic means that the processing effect is due to the nature of the inflections, as also found by Emmorey (1995) in her ASL studies, and not due to e.g. the serial nature of inflections. Other experiments should shed more light on the processing differences in noun inflections.

5.4 Plural dominants

The account of dual route processing holds not only for regular versus irregular plural forms; different processing paths have also been found for other factors. Levelt et al (1999) had an interesting finding. They measured the reaction times in a picture-naming task with plural dominant nouns and singular dominant nouns. Subjects were shown pictures containing one or two identical objects, they were asked to use the singular or plural. The pictures contained objects that are considered to be plural dominant or singular dominant. „*Hands*“ or „*eyes*“ are examples of plural dominants because they occur more frequently in the plural than in the singular form. Such information can be deduced from a count in a representative corpus. A single-dominant word would be „*nose*“, for instance, as it occurs more frequently in the singular than in the plural. Interestingly, the plural forms of either singular or plural dominants evoked longer time latencies than their respective singular counterparts. However, there was a significant difference in the overall reaction times between singular and plural dominants. Plural dominants were about 50 ms slower than singular dominants, in either their singular or plural form. This is a remarkable finding, since plural-dominant plurals had been assumed to exist as separate lexical entries in the lexicon. In a dual-route model of morphological processing, their frequency of occurrence would put them into the category of separate lexical entries, unlikely to be computed by rule application. In case of single lexical entries, no plural suffixation rule would have to be executed and therefore, we would expect their latency to be shorter than for singular-dominant plurals. The results of the experiment have been explained by assuming a competition effect due to co-activation of two lemmas in the mental lexicon. This means that during the process of lexical selection, both the plural-dominant singular form and the plural-dominant plural form are activated. The competition effect between the two lemmas causes the longer latency for plural dominant plurals, similar to an increase in a cross-language lexical decision task when the lemmas of two languages are activated (Holmes, 1996).

5.4.1 Singulars and plurals in Dutch

Baayen, Dijkstra and Schreuder (1997) have carried out reaction time experiments with singulars and plurals of singular dominants and plural dominants respectively. They used Dutch nouns with regular plural inflections as their target items in a timed visual lexical decision task. The independent variables were stem frequency and surface frequency for the noun. The underlying question was whether the parallel dual-route model for inflections holds not only for the regular versus irregular plural paradigm but also for differences in overall stem frequency. The experiment showed that the cumulative frequency of the noun (i.e. the sum of the stem frequency and the respective surface frequency of either singular dominant or plural dominant plural determined in a written corpus of 42 million entries) is a determinant for the reaction times in their singular forms, and the surface frequency is a determinant for the RTs in the plural forms. Thus, frequency can be seen as a determinant for differing parsing routes. Baayen, Dijkstra and Schreuder (1997) developed a mathematical model that exactly predicts the reaction times for various nouns, depending on their stem and surface frequency. Their model turned out to be a reliable predictor for the reaction times observed in the experiments.

Frequent nouns were responded to faster than nouns with a low frequency, suggesting that they have full-form representations in the lexicon. This sheds new light on the question of automatisation. Does this mean that an automated inflectional procedure results in a separate entry in the lexicon since highly frequent items seem to be processed faster than items of low frequency? The experiment was repeated as an auditory lexical decision task for Dutch plurals, including the plural -*en* affix as well as the plural -*s* affix. The intention was to find out whether the changed input modality as well as the extension to different plural affixes would maintain the validity of the test results. Furthermore, the replicated task has been compared to a visual progressive demasking task in order to rule out task-specific effects (Baayen et al, 2003). In this revised experiment, as in the original one, the researchers observed a main effect of stem frequency (high frequency words elicited shorter response latencies), a main effect of number (plurals were reacted to more slowly than singulars) and a main effect of dominance (plural dominants were reacted to more quickly than singular dominants). As before, the researchers noted a surface frequency effect for the plural forms and a stem frequency effect for the singular ones.

The existence of full-form representations of plural forms in the auditory experiment may also be a result of the phonologically altered (i.e. shortened) stem in the plural compared to the longer stem in the singular. In order to rule out a task effect, the outcomes of the lexical decision task were compared to the results of a visual progressive demasking task. The findings are consistent in both tasks, for both modalities and for both plural endings. Therefore the outcomes argue for full-form representations of plural forms in the lexicon (Baayen et al, 2003). Included in the test battery was a comparison of Dutch noun plurals with verb plurals. Contrary to prior visual experiments, Baayen et al. found in this auditory lexical decision task that verb plurals must also have independent representations in the lexicon. The experiments demonstrated that such frequency effects are not specific to either task or modality, or affix; similar results have even been reported for different languages (Baayen et al., 2003).

This series of experiments is not compatible with the idea of full-form representation for irregulars only. A dual-route model (separate processing paths for regular versus irregular inflection) does not predict frequency effects. Hence, the experiments by Baayen et al. suggest that different processing routes do not depend on the type of inflection. „Independent representations of plural forms for both nouns and verbs even for regular affixes argues against the view that storage in the mental lexicon is reserved for irregulars only“ (Baayen et al, 2003: p. 28). As mentioned before, frequency seems to be of importance and processing differences are a matter of degree of frequency. The processing differences between full-form lexical entry and syntactically computed suffix may be gradual. There is no clear-cut divison between the two kinds of processing but rather a gradual spectrum depending on lexically inherent properties such as frequency. Neef (2001) acknowledges that separate lexical entries for irregulars are a matter of degree of irregularity. Thus, separate lexical entries for regulars are also a matter of degree of regularity, depending on lexical frequency.

5.5 Frequency

The distinctions between regular and irregular inflections and between singular and plural dominants are not the only factors influencing processing times. Mere word stem frequency has shown to be a relevant factor in the way inflected and uninflected words are processed in experimental tasks (Baayen et al, 2003). If frequency influences reaction times in processing experiments, and if the degree of frequency matters for processing, it is hard to disentangle all the factors that play a role in processing when measured in the time course of lingusitic processing. Can one therefore still speak of a simple dual processing account if processing is a matter of degree? And if it is a matter of degree, one should also be interested in the watershed between the two extremes. In other words, at what degree of frequency and at what degree of irregularity does the inflected word become a full-form entry in the mental lexicon? It is the question of such a threshold that is interesting in the experiment of this study which addresses the frequency issue in language processing. In the light of language acquisition, it is interesting to know at what point in language development the threshold will be reached and whether this point corresponds with the automatisation of linguistic rules.

Frequency effects, like those found by Baayen, Dijkstra and Schreuder (1997), prove to be problematic for dual-route models of morphological processing where only regularly inflected forms are processed by rule. Dual-route models do not allow a frequency-based processing preference. So far, the focus of morphological research has distinguished between rule computation and rote storage of linguistic forms. Instead, one has to imagine a frequency threshold that determines the processing route. In such a processing model, the frequency of the linguistic item weighs and determines the strength of activation. As a result, the linguistic form may either be processed more quickly or more slowly. In very fast processing, full-form storage may be more likely since the time course does not allow for the application of rule computations.

Again, the question of the interplay between frequency and automatisation is to be considered the determinant factor for the processing route. If frequency contributes to the automatisation of certain morphosyntactic features such as inflections and therefore influences the processing route, it should show in a performance difference for native speakers and non-native speakers whose structures are not automatised. Segalowitz and Segalowitz (1993) propose that automatisation is a matter of restructuring subcomponent processes and not just a speeding-up effect. Such a qualitative change can be explained for the plural inflection that is retrieved as component parts of the lemma in the lexicon. The normal rote computation of the plural inflection (noun + pl) in a plural context becomes automatised when the word naturally appears frequently in the plural in case of plural dominants. Repeated usage of the word in the plural form may lead to routinised components of lemma and inflection in the lexicon; that is the lexical entry already comprises a plural form and competes with the ´old´ rote computation. For example, the plural form of "*hands*" may become a full-form lexical entry due to its plural-dominant nature and may be based on automatic memory-based retrieval, as Logan (1988) calls it in a context exceeding the realm of linguistic processing. Then, it may compete with the computed form. Due to increased activation levels, one may be more successful than the other which will then be suppressed.

Jennifer Hay (2001) has found that it is not mere frequency that determines morphological decomposition, at least for derivational morphemes. She claims that highly frequent forms are not highly decomposable nor are they very semantically transparent. A perfect example would be the form "*clothes*" with a highly idiosyncratic meaning which cannot be decomposed into its inflection and the singular form. The way a word is accessed, Hay claims, should affect the mode of representation and also its semantic transparency. In fact, the link between frequency and opacity is not new, explaining the apparent cost of efficiency for full form representations since they present a substantial memory load if they have to be stored in the mental lexicon. Therefore, all full-form representations should have a kind of evolutional advantage over their decomposable counterparts which makes this memory cost effective. In most cases, this evolutional advantage is the semantic idiosyncrasy or opacity.

Hay (2001) looked at derived forms of different frequencies, and compared their definitions in the dictionary. She found that those words that are relatively more frequent than their base, thus displaying a high relative frequency, seem to be less semantically transparent, judging from the number of definitions in the dictionary. In her experiment on relative frequency, Hay showed her subjects a randomised list of prefixed and suffixed word pairs (differing in frequency), asking them to judge their degree of decomposability. Her experiment showed that the frequency of the base form is an indicator of its assumed decomposability. However, the frequency of the base in relation to the frequency of its derivation is the crucial determinant here. Hay calls this the relative frequency (2001). She points out that the nature of the affix is an important factor in decomposability since the temporal nature of speech distinguishes between prefixed and affixed derivations. Prefixed words are predicted to be generally less decomposed than suffixed forms. Hay´s data actually confirm this assumption (2001). For methodological purposes, the conclusion to be drawn from this study is the importance of relative frequency. In experiments involving frequency, the researchers must control for the frequencies relative to one another (Hay, 2001). This will also be considered in the experiment in this study, even though only plural inflections will be examined and not, as in Hay´s experiment, derivational affixes.

Likewise, Sereno and Jongman (1997) have reported frequency effects for inflected nouns and verbs. They set out to review the existing findings in the literature of relevant factors influencing the word recognition process. They found that grammatical class membership had hardly been studied; therefore they devised a lexical decision task to observe possible reaction time differences between nouns and verbs. Such an approach is important because, so far, most studies have concentrated on either inflected nouns or inflected verbs. In an earlier section, I showed that it is impossible to simply consider that the findings on inflectional processing routes of the priming task of verbs by Münte et al (1999) are valid for nouns. An experiment such as the simple comparison between reaction times of nouns and verbs as designed by Sereno and Jongman (1997) should show whether grammatical class actually matters for the processing procedure. The stimulus items were selected and matched for total frequency (sum of the frequency of inflected and base form). In this case, the response latencies for verbs were significantly longer than the response latencies for nouns. These processing differences (in this case the differences are of a quantitative nature) were ascribed to frequency effects since the overt frequency of uninflected nouns is much higher than that of uninflected verbs. If this is true, differences of inflectional structure between nouns and verbs are determinant, since both were matched on the basis of total frequency.

The second experiment was to test this hypothesis. Only nouns were used as stimulus words and the two conditions were nouns with relatively high-frequency singular forms and correspondigly low-frequency plural forms versus nouns with relatively low-frequency singular forms and correspondingly high-frequency plural forms. The overall frequency, however, was the same for both conditions. If it is inflectional structure that determines the processing route, there should be a processing difference between the two conditions. If total frequency matters, no difference should be found since the data is being matched for sum frequency. In the first run, subjects were shown the base forms only. In the second run they were shown inflected forms only. For the pure base forms, the experiment revealed that response latencies for high-base frequency nouns were lower than those for low-base frequency nouns. In the second run, where only pluralised nouns were presented, subjects‘ response latencies were higher for high base frequency/low plural frequency forms than for their reverse counterparts (low base frequency/high plural frequency). The condition with inflected nouns showed that the surface plural frequency determines the speed-up effects, not the base frequency of the word form. With other words, the frequency of the presented form appears to matter.

The third experiment in this series was set up to measure the effect of overall frequency. Experiments one and two both matched the stimuli for equal overall frequency and therefore only showed whether base or inflected frequency is dominant. In the third experiment, base form frequencies remained stable, and the differences in plural frequencies determined the overall frequency. If total frequency matters, there should be a difference between conditions for the singular nouns, even though in both conditions the base frequency is the same. Again, the first experimental run only presented singular forms and the second one only plural forms. The first run (only singulars) revealed no effect for overall frequency. In the experimental run with only pluralised nouns, the RTs were faster for equal-base/high plural frequency than for equal base/low frequency nouns. The third experiment clearly shows that surface frequency for inflected forms determines processing speed.

Both the experiments by Sereno and Jongman (1997) as well as the data from Hay (2001) cast doubt on the dual-route model. A strict separation of processing routines between regularly and irregularly inflected words is hard to sustain when even regularly inflected words may behave like separate lexical entries depending on their surface frequency. Rather, the studies provide evidence for a unitary system that is highly responsive to surface frequency effects for each form, even for such productive morphology as pluralised nouns. All further research on inflections has to account for frequency effects. What can definitely be concluded is a frequency effect for both verbs and nouns and base- as well as inflected forms. It is the aim of my experiments to find out at what point those frequency effects can be detected in second language development. In terms of theoretical processing models, the question remains whether, at some point after they exceed the frequency threshold, highly frequent words enter the lexicon as separate entries. Or, should the idea of a dual route model be abandoned completely and the division in lexicon and grammar be assumed an artificial one?

Hypothesis: Lexical access of regular plural forms may rely on different processing procedures, partly depending on lexical frequency. Second language learners rely on the frequency of input to build their lexicon. If frequency should become apparent at a given developmental stage, this should show in the learners´ processing times before and after this point in time. In case of a unitary processing system a gradual change in processing should be observed in learners of differing developmental stages.

The next chapter will introduce the lexical decision test paradigm and will describe the first experiment used in this study. Factors such as frequency and dominance shown to have an influence on morphological processing will be considered in the test design.

6 Methodology

The previous chapter described morphological processing in detail. This chapter will introduce the lexcial decision task paradigm as well as the phrasal grammatical judgment task paradigm used in this study. Section 6.1 also reviews the most important psycholinguistic findings observed in the lexical decision task relevant for the test design in this study. Although touched upon in a more general context in section 3.2 and 5.4.1, the main effects specific for lexical decision experiments will be summarised here. A detailed account of the methodological considerations follows. Section 6.2 focusses on the phrasal grammatical judgment task, introducing the design and the exact procedure of the task and explaining how it differs from the first experiment.

6.1 Experiment I: Lexical decision

6.1.1 What is a lexical decision task?

In order to investigate linguistic processes during the actual time of processing, a multitude of experimental methods can be used, including visual or auditory priming experiments, neuroimaging techniques and gating experiments. The lexical decision (hence abbreviated as LD) test paradigm is one of the experiments that records the subjects´ reaction time to a linguistic event while the stimulus is ongoing. One of the first reaction time tasks using the lexical decision paradigm was the frequently cited Meyer and Schvaneveldt (1971) LD task. Meyer and Schvaneveldt used a lexical decision task to determine the strength of semantic word associations when a pair of words was presented simultaneously and judgements had to be made about the “sameness” as well as about the nature of the word (real or nonce). Meyer and Schvanelveldt also built on previous tests in the late 60´s and early 70´s that used the lexical decision paradigm.

While responding to the linguistic stimulus, the participant is not aware of the actual purpose of the experiment. On-line experiments allow the researcher to draw conclusions about the mental processes involved in language task. Therefore they “provide insights into the relationship between grammar and processing” (Marinis, 2003: p.158) because the researcher can distinguish between off-line grammatical surface competence and its underlying processing procedures[7]. A learner might have mastered a grammatical structure in language comprehension or production but might process it differently from a native speaker. This new on-line methodology has become an important tool to investigate processing procedures that otherwise would remain invisible. On-line tasks may therefore pinpoint the exact nature of learners´ language processing and ultimately lead to an understanding of how native-like processing can be learned.

In the LD task, it is the reaction time of the word recognition process that is being measured. A lexical decision task measures the speed of lexical access or lexical retrieval, that is the time between reading or listening to a word and comprehending its meaning. The LD task used in this study measures the subject´s reaction time to a string of letters represented visually on a screen. The subject´s only task is to decide whether the string of letters constitutes a real word or a nonce word. Section 6.1.3 will introduce the method and procedure in detail. In previous experiments, as reported below, lexical decision tasks have shown stable effects for different linguistic phenomena.

7. The kind of competence referred to here is in no way related to grammatical competence in a generative sense. Here, grammatical competence refers to the learner´s knowledge of grammatical structures in off-line tasks (production or comprehension).

6.1.2 Robust findings

6.1.2.1 Word length

Researchers have found conflicting results on the question of whether longer words take more time to process than shorter words. Some of the psycholinguistic findings have already been described in section 3.2 but those relevant for lexical decision tasks will be summarised here. Reviewing several lexical access models and the major experimental findings of psycholinguistic experiments in chapter 5 of his book, Garman (1990) summarises previous findings on the word length effect in visual recognition tasks and states that there is virtually no convincing word length effect in recognition tasks, as Forster (1976) discovered earlier. This means that in the visual lexical decision tasks reviewed here, subjects do not require more processing time for longer words than for short words.

These findings might suggest a parallel processing mode for visual lexcial recognition with whole word forms processed at once, versus a serial processing (letter by letter) mode for auditory word comprehension. Parallel, here, means the simultaneous processing of all letters at once. If the processing times for long and short words are similar, the letters cannot be the decisive unit in visual word perception. Visual word recognition times are not accumulated recognition time for letters or graphemes (written letters). Rather, the whole word might be recognised as a complete unit, and thus conform to the biological notion of holistic pattern recognition, as described by Garman (1990, p. 208/209):

> Two striking properties of writing systems, as compared with natural objects, are that they are typically composed of just two-dimensional elements (rather than three), and have a highly conventional structure. Letters of the alphabet do not (like tables) have objects lying on top of them, obscuring or extending their outlines; nor do they move about (like animals), or change their size and shape as you look at them (as balloons do while they are being inflated). They consist of a limited number of light-dark contrasting patterns, in terms of vertical, horizontal, slant, straight and curved edges, in various, but limited permutations. As such, writing systems constitute visual arrays that avoid many of the really problematic issues in our understanding of visual-shape perception. It is tempting to conclude from this that they are in some sense easy to process, a convenient feature for a system whose task is to represent the complexity of language for the reader or writer. But it is not clear what sort of early processing is involved, since it is a matter of debate as to what sorts of properties of the array the visual system makes use of.

Garman goes on to present studies both in favour of and against the assumption that the grapheme is the mediating unit in whole-word recognition. What is of interest here is only the fact that visual lexical decision tasks do not show any word length effect when graphemes are the processing units.

The sequential nature of spoken word recognition, however, has been supported by Grosjean (1980) who carried out gating experiments and found that subjects recognised words reliably before their offset (the end of the spoken word). Similarly, van der Lugt (1999) found that in discriminating words from a string of sounds the likelihood of the word-initial phoneme sequence is much more important than the likelihood of the phoneme sequence of a word´s offset. This means that during auditory processing different points of a spoken word have different importance for word discrimination and a spoken word may not be processed as a complete unit. It seems natural to conclude that spoken word recognition has a different nature from visual word recognition and operates with different perceptual units. The spoken word percepts are phonological segments, cued by prosody, phoneme assimilation and other suprasegmental cues. As such, spoken words have a "uniqueness point" (Marslen-Wilson, 1984) as described in 3.2.4. and can be recognised before the whole word is spoken.

Despite the lack of evidence (Garman, 1990) for a word length effect in visually presented words in a lexical decision task, the visual LD tasks in this study nevertheless use words similar in length. All target words are monosyllabic. Syllables are phonological units and therefore not comparable with visual perceptual units, such as graphemes or letters. In using the same syllable length for all stimulus words, though, the task design accounts for possible grapho-phonic perceptual representations of words. These are thought to be the abstract perceptual units that incorporate auditory and visual features of a word (Garman, 1990) and might play a role in processing.

6.1.2.2 Real word effects

One of the robust findings in LD tasks is the reaction time difference between real and nonce words. Phonologically legal nonce words take longer to be identified than real words (Garman, 1990). It is important to note that the nonce words are phonologically legal because nonce words might otherwise be rejected right away due to impossible phonological rules. Such findings in a visual task are consistent with the "spoken word recognition" or cohort model of Marslen-Wilson and Welsh (1978) described in section 3.2.4 which assumes a stepwise recognition of word segments, depending on the number of competitors beginning with the same word segment (the cohort of possible word selections).

Even though it is a robust finding, the real-word effect is not relevant in the experimental design of this study. The nonce words in this study are all distractors and only the 70 real target words out of almost 400 stimulus words really matter for the findings of this study (cf. section 6.1.7.3). Nevertheless, such a robust effect should show even if it is not the focus of the study. If it is also found in the subject groups presented here, it is perfectly legitimate to assume that the experiment is a representative one and reproduces previous findings. The subjects´ behaviour regarding the real-word effect will be presented later together with all results.

6.1.2.3 Frequency effects

Another effect found consistently in LD experiments is the word frequency effect (Forster, 1976). As Garman (1990) points out in his overview of various psycholinguistic findings, high-frequency words are processed much faster than low-frequency words; this effect is repeatedly supported as reported in section 5.5. Interestingly, neuroimaging techniques such as event-related potentials (ERP) and magnetoencephalography (MEG) studies have also revealed that semantic processing is invariably related to probabilistic factors such as frequency or repetition. The N400 amplitude - negative brain activity occurring 400 ms after presentation of the linguistic stimulus and associated with semantic processing - has been shown to be affected by factors such as frequency, repetition and semantic association (Marinkovic, 2004). The need to incorporate probabilistic factors into models of language processing has also been acknowledged by scholars such as van der Lugt (1999), Baayen (2003), Manning (2003) and Bod et al (2003).

6.1.2.4 Semantic effects

Another relevant factor that leads to noticeable results in primed LD tasks is the semantic priming effect. Semantically associated words may prime each other in a primed LD task (Garman, 1990). That is, they may influence each other´s RTs. The word "thunder" presented as a prime stimulus shortly before the target word "lightning" reduces the reaction time of “lightning” if presented at the right time interval (Levelt, 1989) as compared to the unprimed base condition. In some cases, primes may also increase the response latency. This shows that in spite of time pressure and the irrelevance of the semantic nature for the task in question, the meaning of a word does play a role and cannot be ignored in real-time processing experiments such as lexical decision tasks.

A semantic component influencing processing speed will also be studied in the second experiment of this study (see section 6.2). In order to ensure that the basically semantically unrelated stimulus items do not accidentally prime each other in the LD experiments of this study, the order of presentation of the stimulus items changes randomly from experiment to experiment.

6.1.3 Method

Before the experiment was carried out under real conditions, the design and the set up of the task were tested in a pilot study. The pilot study was needed to fine-tune the experiment and to detect possible fallacies in the experimental set-up. After the pilot study, some words were exchanged because a couple of participants were not sure about two of the nonce words. Although the nonce words were not considered target items and merely served as fillers, they needed to be changed in order to match the feedback given by the programme. Another amendment was changing the timing at which the feedback ended after the trial items. The overall setup and all the stimulus words, however, remained the same in the real trial. The results of the pilot study will not be reported because the experimental conditions were not perfect and are not comparable to the final experiment.

6.1.4 Participants

There were two sets of participants, one for the pilot study and one for the real test condition.

6.1.4.1 Pilot Study

The participants in the pilot study are 47 near-native speakers of English and 14 native speakers of English. The near-native speakers are German students reading English; most are in their final years at Paderborn University. Some of them had spent a semester in an English-speaking country while others were close to finishing their studies. The native speakers were enrolled as exchange students at Paderborn University. The participants were recruited in one of the two ways: by addressing them directly in class or by sending an email with a request to participate in a linguistic experiment. An address list of the foreign students was provided by the foreign office. The students took part on a voluntary basis and were not rewarded for participating.

6.1.4.2 Real test condition

In the real experiment, two main groups of subjects participated: native speakers of English and learners of English. These two main groups were subdivided into different age groups.

The native speakers were divided into two experimental groups:

- **29 native speakers** aged 16 and 17

- **34 native speakers** aged 18 to 39

Native speakers are necessary in order to obtain the native speaker standard for frequency and dominance effects, as discussed above. Only in the light of the native speaker results does it become possible to analyse the learner data.

Of the 63 native speakers of English tested, 29 were 16 / 17-year-old secondary school pupils in their A-level year at the King´s School Gütersloh. These students are all children from families connected to the British army stationed in or near Gütersloh, Germany. Usually, these British pupils have little contact with Germans and grow up in an almost completely monolingual environment. The remaining 34 native speakers are British soldiers from the Electronic Unit at the Barker Barracks in Paderborn, Germany. The ages of these soldiers ranged from 18 to 39. The main reason for testing two groups of native speakers is the fact that the group of school pupils is of the same age as one of the language learner groups described below. The soldier group includes only adult native speakers usually considered the linguistic norm. If an effect can be observed in a learner group and also in one of the native speaker groups, it cannot be specified as a typical learner effect but may simply be an age effect. The two groups of native speakers are therefore necessary in order assume correct findings.

The participants were assigned to take the test as part of their regular duties. The pupils had to take it during one class hour whereas the soldiers were ordered to participate by their superiors. The groups were not rewarded; the only incentives were chocolate bars for all of them. The soldiers especially enjoyed this change from their regular activities and were rather competitive. My visits at the barracks and at the British army school were very much welcomed by the school board and the liasion officer, whom I contacted beforehand. The British army is generally keen on keeping up good relations with local Germans and was therefore pleased to have a visitor from the local university. Some of the soldiers were even sent over to the university building one morning in order to take the test at the computer lab as they had to march to the university and run back, the test was incorporated into their sports day.

Six German learner groups of gymnasium pupils participated in the experiment.

Table 3. *Six learner groups participating in the experiment*

number of participants	***school year***	***average age***
27	*7*	*13*
24	*8*	*14*
39	*9*	*15*
20	*10*	*16*
17	*11*	*17*
18	*12*	*18*

In order to capture different age groups with different language skills, I tested students in six different grades from year 7 (with three years of English language instruction) through year 12 (with eight years of English language instruction). All these classes started with English language instruction in their first year of secondary school (class 5). I decided to differentiate between school years when looking at development in processing because frequency and dominance effects are closely related to input and it is the input that crucially differentiates the learner groups at school. The data were collected at the end of the school year at the Reismann Gymnasium in Paderborn and the Gymnasium in Delbrück.

Like the native-speaker groups, the administrators at both schools and all the school teachers were very co-operative and helpful in carrying out the experiment. Again, the students were not rewarded except with chocolate bars.

6.1.5 Procedure

In the real experiment, all groups worked in the IT room of either the school or the University. Each group took the test collectively during a lesson hour. The pilot test was carried out on an individual basis in an office at the university. Each experiment took about forty minutes. The subjects sat in front of the computer screen and were asked to follow the instructions on the screen (see appendix).

Before they started, participants were informed that they would be presented with ten test items before the start of the actual experiment. After responding to the set of test items, participants were allowed some time to relax and ask further questions before the real experiment started. After pressing the space bar to request the actual experiment, the subject saw a string of letters in the middle of the screen prompted by a fixation point preceding it at the same position. The subject then was to decide as quickly and accurately as possible whether or not the presented string of letters was an existing English word. The computer recorded the subject´s reaction time for each stimulus from the onset of its presentation. The items were presented randomly in a different order at each test run. After each response, the participant immediately received feedback telling him whether his answer was correct. The feedback also provided information on the exact reaction time in milliseconds and the percentage of correct answers for the experiment at that point.

6.1.6 Apparatus

Setting up the lexical decision experiment required a software package that facilitated the programming of all variables in a user-friendly way. The software package used to program the experiment is a psycholinguistic research environment called E-Prime (http:// www.pstnet.com/e-prime), commercial software developed on the basis of an earlier program called PsyScope. Psyscope was the first such program developed at Carnegie Mellon University for use in psychological experiments. Its successor, EPrime, was described by Marinis (2003: p.157) as "the most user-friendly experimental software for PCs within the Windows environment if the department does not have technical staff to support the experiment". It was chosen mainly for these reasons.

E-Prime contains various domains for specific tasks. The programming domain is called E-Studio and is used to develop an experiment. E-Run is the domain used for running the experiment in a test situation and E-DataAid helps to manage the data obtained after running the test. The data files are large files of roughly about 40 A4 pages per subject. If the test yields some irregularities or unusual results, E-DataAid makes it possible to check each individual variable, line by line. These lines contain the biographical information of the subject, the response time, the nature of the response, the response accuracy and the stimulus item features (frequency, dominance, etc.) for each stimulus word. Another domain is E-Merge which merges separate data files if subjects of one group are recorded in different test runs but must be included in the group. With all these domains and their specific functions, the software is easy to understand. A helpful graphic interface allows the user to develop his or her own experiments without prior experience or help from technical staff; a helpdesk can be contacted either online or by phone. The manual included in the software package contains an introduction to programming an experiment from scratch and contains sample LD experiments with simple variables. E-Prime was therefore chosen as the most practical software for the purpose of the experiments in this study.

The experiment was run on regular PCs with an Intel Pentium III processor and 1024x768 pixel, 17-inch monitors positioned about 50 cm away from the subjects. Reactions were registered on a keyboard which fed into the E-Prime reaction time recording device. The stimuli were presented in a bold 18-point black Courier New font in the centre of the screen. No additional software was needed; EPrime provided the core of the LD task.

6.1.7 Test design

In RT experiments, participants have to react under time pressure; thus the LD task is considered an on-line task. The word recognition task is merely an excuse to elicit real-time recognition effects for factors hidden in the design of the task. Participants are informed that the purpose of the experiment is to decide correctly between real and nonce words. This statement obscures the researcher´s genuine intention in such on-line experiments. On the basis of the results, the researcher can then compare reaction times for various factors that are part of the word-recognition process. He or she can, for instance, compare the reaction time between singular and plural word forms or the RT between nouns and verbs. The subject, however, has no information about such crucial test variables. LD tasks are therefore an indirect measurement of real-time linguistic processes not subject to the participant´s control. The considerations involved in selecting the stimulus material will be presented in the following sub-sections.

6.1.7.1 Mode of linguistic input

In LD tasks, the participant reacts to visual or auditory stimuli and the task always measures word recognition, hence lexical decision. The LD task presented here is a visual one; the stimulus words are presented in written form on a computer screen. Moreover, this experiment was set up to study the processing nature of inflectional plural endings. Whether or not the mode of presentation of the linguistic stimulus influences the quality of linguistic processing can be debated from various points of view and was touched upon in section 6.1.2.1, in discussing the word-length effect.

In her overview of the latest neuroimaging techniques for language processing, Marinkovic (2004) describes the pathways of word processing in terms of spatiotemporal brain activity for both modes of presentation, visual and auditory. She demonstrates that both modes of input eventually end up in a supramodal processing network in a time sequence of just about 230 ms after stimulus onset, independent of input modality. Accroding to this model, the time course for general processing is therefore fixed, regardless of input mode. Thus, an N400 amplitude found in neuroimaging experiments and associated with semantic processing occurs 400 ms after stimulus presentation, be it visual or auditory. Her findings suggest that input mode is not relevant for reaction time experiments.

Research findings (e.g. Baayen et al, 1997) suggest that for subjects to recognise words with inflectional endings what is important is the time needed to compute the word form from its morphological components, more than the mode of representation. Because a word may be represented as a phonological unit and as a visual unit in linguistic memory, the mode of input may play a role for linguistic memory. Crucial for inflectional processing, however, is the way a word is morphologically composed (Baayen et al., 1997). Moreover, in her dissertation Poelmans (2003) demonstrated that the differences in reaction time for the visual versus the auditory LD tasks for native and non-native speakers were not significant. The significant difference was between the accuracy rates (percentages of correct responses) for the two modalities. Thus, there does seem to be a modality-specific effect for the word recognition process, but only for accuracy, not for reaction times.

I chose to use the visual LD paradigm for these experiements partly for practical reasons, because E-Prime does not allow auditory stimuli in its current edition. Another important reason was explained in section 6.1.2.1: there is evidence for parallel processing of written words as opposed to sequential processing of spoken words. For inflectional processing this means that the subject may decide to classify a string of sounds as a real word as soon as the word stem is completed, even before the subject hears the plural inflection. Therefore, an auditory LD task would measure word recognition but not recognition of the inflection (Grosjean, 1980). In the visual mode, the participant has to classify a string of letters as a complete representation including the plural inflection (Garman, 1990).

6.1.7.2 Selecting the stimulus material

The stimulus words chosen for this experiment had to be carefully matched for frequency and word length (cf. 6.1.2.1 and 6.1.2.3). In order to determine frequency, I took the stimulus words from the CELEX language data base (Baayen, Piepenbrock & Gulikers, 1995) which includes different native language corpora for three languages. Its lexical entries are all taken from a huge variety of written and also some spoken material, representative of all areas of modern life. The English language database contains 160,000 word forms. These word forms are annotated for their grammatical function, e.g. noun or verb, and their lexical characteristics, for instance frequency of occurrence. The database contains all this information for each entry. Sometimes a lemma, such as “bike”, may be a noun as well as a verb. In this case the word form is listed twice in the corpus, once under the heading *noun* and once under the heading *verb*. It should be noted that in this context *word form* refers to a specific form, whereas *lemma* refers to the abstraction of a word. For instance, the word “speak” and the word “speaks” are two word forms for just one lemma [SPEAK]. For English lemma homographs, such as “bank”, the words have been disambiguated on the basis of the 17.9 million word English Collins/COBUILD text corpus

(CELEX homepage, 2001). My LD experiments, however, specifically excluded homographs in order to prevent ambiguous frequency information. The appendix contains a list of all stimulus items.

6.1.7.2.1 Selecting frequency ranges

The entries in the CELEX database are annotated for their frequency of occurrence within the corpus. This is a rough estimate of the relative frequency of the words in the target language. CELEX provides information about the absolute or raw frequency in the corpus: the frequency of each entry calculated per million words. Additionally, CELEX provides the logarithmic (log) frequency, which is used to account for human perception of frequencies. Cognitive psychologists have demonstrated that the same difference between two numbers is perceived differently depending on how large the numbers are (Whalen et al, 1999). Humans perceive the difference between the numbers 7 and 9 to be far more striking than the same mathematical difference between the numbers 5977 and 5979; the perceived difference of the latter pair is relatively less relevant than that of the first pair. Likewise, timed experiments have shown that people discriminate the difference in the first pair more quickly than that in the second pair (Moyer & Landauer, 1967). The logarithmic frequency corrects this perceptive difference and uses a logarithmic formula to transform the absolute frequency into the logarithmic frequency. The linear scale of the raw frequencies is transformed into an exponential one which is much more in line with human perception (Stowe & Kaan, 2001).

I used this so-called log-frequency to select the word stimuli in the described experiments. The log-frequency of the noun list had a range of 0,6021 for the word *writ* and 3,2858 for the word *ones*, the most frequent word in the noun corpus. In absolute frequency, this would be 63 entries for the word *writs* and 34640 entries for the word *ones*. Therefore the log-frequency that accouts for human perceptional irregularities has a range from roughly 0 to 3. Since human perception has already been accounted for, it was possible to simply divide this range into three equal parts for the experiments in this study. The three equal frequency ranges were used to determine low (0-1), mid (1-2) and high (2-3) frequency words in the CELEX databse.

6.1.7.2.2 Determining singular or plural dominance

In addition to the frequency information, the relationship between singular and plural frequency is relevant in this study (cf. section 5.4 on dominance). This relationship is also called the dominance value of a word. In order to study the relevance of the dominance factor in noun processing, the stimulus words needed to be selected for either singular or plural dominance; I did not consider words with equal frequencies in their singular and plural form.

The dominance value of a noun can be determined by comparing its singular and plural frequencies. Therefore, it was necessary to view the annotated words of the CELEX database in such a way that both singular and plural frequencies of a lemma were visible simultaneously. For this purpose, a colleague created a software programme that connected the frequency file of the CELEX database with the wordform file of the CELEX database. The frequency file provides information about word frequencies whereas the wordform file provides information about morphological wordform, i.e. whether the word in the corpus is an inflected plural noun or a verb marked for agreement. Both cases may take one and the same word form as, for instance, in “bikes”. After combining the two files, I extracted all the nouns because all other entries (e.g. verbs, adjectives) were not needed. This facilitated the search for singular and plural frequencies. The new file containing all the annotated nouns was split into a list of all singular nouns with their respective frequencies and a list of all plural nouns with their respective frequencies. Of course each noun would occur in the list as both a singular entry and a plural entry. After creating two lists with the same nouns in either their singular or plural form, I had to fuse the lists in order to view the singular and plural frequency of a given noun at the same time. A joint list was created by sorting the word forms according to their lemma ID number. Each noun was now listed according to its lemma ID number, providing both its respective singular and respective plural frequency. The list could then be exported to Excel, which facilitated the search for relevant stimuli. This process made it possible to create lists with nine different conditions (see Table 4). The variables used in this experiment were frequency (high, mid and low), dominance (plural dominant and singular dominant) and number (either singular or plural).

Table 4. *Nine different stimulus groups*

group	stem frequency	plural frequency	dominance
1	high	high	equal
2	high	mid	singular
3	high	low	high singular
4	mid	high	plural
5	mid	mid	equal
6	mid	low	singular
(7	low	high	high plural)
8	low	mid	plural
9	low	low	equal

Group 7 yielded no entries and was therefore discarded as irrelevant to the search for items. For each group, all the one-syllable words were selected in order to minimise variation in word length. Even though syllables are phonetic units, not orthographic units, I chose them as the relevant measure for word length (as argued for in section 6.1.2.1). After all, reading a visually presented word calls upon a phonological representation of a word. Therefore not only the mere length in terms of written letters but also the phonological segments play a role in perceiving word length. Besides, syllables are usually similar in orthographic length, at least on average.

All the one-syllable words were then checked for their semantic opacity, and, as mentioned above, ambiguous forms were removed.

6.1.7.2.3 Relative dominance

The remaining eight groups were transformed into ten groups with different relative dominance levels. Thus a new variable, *relativity*, evolved because the equal domiance groups rarely have entries with exactly equal frequencies for singular and plural form. Even within the equal dominance group, the frequencies for singulars and plurals vary substantially. The new groups were created by splitting the equal dominance groups into either singular-dominant or plural-dominant groups. The following example illustrates this process. The noun “booth” and its plural form “booths” both appear within the low log-frequency range of [0-1] (cf. section 6.1.7.2.1). The singular form has a log frequency of

0,95 and the plural form has a log frequency of 0,3. Even though they are both within the same low frequency range, the noun is still singular-dominant because it occurs more often in the singular. Since the singular form does not exceed the log-frequency range, I defined it as a singular dominant of low relativity. When I split the equal-dominance groups into the groups of relatively low singular dominant or relatively low plural dominant, I had to differentiate them from other nouns where the singular frequency exceeds the frequency range of the plural or vice versa. I defined such cases as either singular-dominant or plural-dominant with HIGH relativity. Another example illustrates this case: The word ”point” occurs with a log-frequency of 2,57 (frequency range [2-3]) and the plural form occurs with a log-frequency of 1,8 (frequency range [1-2]). The word “point” is therefore singular-dominant with high relativity because the singular frequency is in another range than the plural frequency.

The equal dominance groups are those that comprise both singular- and plural-dominant words of the same defined log-frequency level, for instance in the range of [1-2]. However, the degree of singular or plural dominance was to play a role. I split these equal dominance groups into either low sg or low pl dominance, as determined by the frequency value. The low dominance label now denoted dominance of the same log-frequency interval (e.g. *low plural dominant* plural with a frequency of 1,6 and stem frequency of 1,4). The high dominance label now denoted dominance of a higher log frequency range (e.g. *high plural dominant* plural of 2.6 and stem frequency of 1,4). Since the stem and the respective plural both lie within the same frequency interval, I wanted to prevent the occurrence of almost equal dominance levels. Therefore I chose log frequency values that contrasted as much as possible within the respective frequency interval. The final groups, as defined, are presented in Table 5.

Table 5. *the stimulus groups re-sorted*

group	stem frequency	dominance	relative dominance
1	high	plural	low
2	high	singular	high
3	mid	singular	low
4	high	singular	low

Table 5. the stimulus groups re-sorted

group	stem frequency	dominance	relative dominance
5	mid	plural	low
6	mid	singular	high
7	low	plural	high
8	mid	singular	low
9	low	plural	low
10	low	singular	low

6.1.7.3 Number of target words

I then selected stimulus words for each group. Each group contained six to eight words, with an overall average of 7 words for each of the ten groups. These 70 stimulus words or 140 forms (singular and plural of the corresponding word) and their singular and plural forms were distributed evenly into one of the two versions of the same lexical decision tasks in order to prevent a singular and its corresponding plural from occurring in the same experiment. Thus, there were two LD experiment versions, with each containing 70 stimulus words. Each experiment also contained 120 filler words. These filler words were randomly taken from the CELEX database. They were mostly one-syllable words in their uninflected form. Some, however, were bisyllabic and inflected, for instance “called” or “deadly”. The filler words were selected for various word classes in order to shadow the noun targets of the stimulus material and were therefore used as distractors. In addition to the real words, I created an equal number of pseudowords (nonce words) in order to ensure an equal LD design with the same number of real and nonce words. Nonce words were created by exchanging vowels or letters from the stimulus and distractor words. These nonce words have different recognition points; that is, some became nonce words word-initially but I also ensured that most became nonce words only at word offset. The reader then has to read the complete word before actually being able to judge the word as a nonce word. Before the actual experiment, 10 practice items were included. Altogether, in either version of the experiment, the subjects had to deal with 390 items:

- 70 stimulus words (singular or plural nouns);

- 120 filler words (existing words of other word classes);

- 190 pseudowords created from the stimulus and filler items (one syllable nonce words); and
- 10 practice items (randomly taken from all of the above categories)

Two versions of the LD task were used in an alternate way, to ensure that each stimulus word was presented as often in its singular as in its plural form. The same word did not appear in the same test version in both its singular and its plural form.

6.2 Experiment II: Phrasal grammatical judgement

6.2.1 Rationale

The purpose of this phrasal grammatical judgement task is finding out whether a word´s meaning on a conceptual level influences processing quality. In the case of plural processing, this would involve the noun´s *numerosity.* The factor *numerosity* is independent from *frequency* and *dominance*. It simply denotes the kind of number information carried by the plural noun. Some aspects of number information were presented in Chapter 4, which focussed on the various meanings a plural may have.

In this particular experiment, I want to compare *conceptual duals* with *conceptual plurals*. A dual noun, as explained in Chapter 4, is a plural noun that denotes a paired entity, like “wings” or “boots”. The group of *conceptual plurals* includes all nouns that usually refer to more than just two, ideally to a large quantity of something. C*onceptual plurals* means nouns like “tears” or “miles”, or “words”. The experiment is designed to test whether this conceptual information has any influence on the processing quality. The English language encodes both dual and plural information in the same plural morpheme; there is no formal distinction between the two. Yet, conceptually, dual nouns are mostly paired entities; they do not indicate a large quantity of something. Seldom does the word “wings” without a determiner make us think of ten or twenty wings; instead, we would immediately think of a pair of wings.

I matched the plural and dual nouns (using only the pluralised forms) with two types of numeral determiners. In the first case, I used the numeral “two”. Then, I used monosyllabic numerals, comparable to “two”, that refer to a number larger than two, e.g. ”ten”, “five” or “eight”. This resulted in four different kinds of noun phrases:

- a dual determiner + dual noun (as in: *two eyes*)

- a plural determiner + dual noun (as in: *ten eyes*)

- a dual determiner + plural noun (as in: *two tears*)

- a plural determiner + plural noun (as in: *ten tears*)

For the dual nouns, I expected faster reaction times for the noun phrases with a dual determiner than for noun phrases with a plural determiner. I did not expect the plural nouns to be influenced by either dual or plural determiner. Most of the plural nouns referred to a large quantity, certainly larger than either “two”, “five”, “eight” or “ten”. Therefore, processing should not be influenced by the numerosity of the determiner.

If dual nouns in a noun phrase evoke different reaction times depending on the determiner in question, then the conceptual information of the noun´s numerosity influences processing speed. Even though the subjects were asked to judge the noun phrase on grammatical acceptability, not on numerosity, the conceptual information might have influenced their decisions.

6.2.2 Method

The phrasal grammatical judgement task was similar to the lexical decision task and worked in exactly the same way. However, the focus of the task was not in deciding whether the presented string of letters constituted a word. Instead, participants now had to decide whether or not the presented phrases were grammatically correct. They made a decision and pressed the appropriate response key and the computer recorded their reaction times.

6.2.3 Participants

The groups of participants were almost the same as for the previous lexical decision task described in section 6.1.4. During my visit to the schools and the barracks, the subjects first took the lexical decision test and after a little break the phrasal grammatical judgement test. For two of the Gymnasium classes, I was not able to collect the data for both tests because the pupils had to rush to a different lesson. I therefore lack data on the 8th year class and one of the two classes in year 9. Some individual pupils who took the lexical decision test were excused from the second test because they took too long on the first one or had to leave early.

Altogether, 169 subjects participated in the phrasal grammatical judgement test: 63 native speakers and 106 German learners of English from school years 7, 9, 10, 11 and 12.

6.2.4 Procedure

As in Experiment I, I used the IT room at either the school or the University to administer the experiment. Each group took the test collectively during a lesson hour. This experiment took about 15 minutes. The subjects sat in front of the computer screen and followed the instructions as presented on the screen (see appendix).

The response keys on the keyboard were marked with red and green tape in order to facilitate quick responses. The subjects were asked to use both hands and keep their two index fingers on top of the keys during the whole experiment in order to respond as quickly as possible. The key on the right-hand side was chosen to be the response key for the correct words because most participants are right handed and would therefore have a right-hand advantage. Since all the target stimulus phrases were correct, the subjects should have responded to all of them with the green right-hand key for “correct phrase”.

After responding to the set of practice items, participants were allowed some time to relax and ask further questions before the real experiment started. After pressing the space bar to request the actual experiment, the subject saw two words in the middle of the screen prompted by a fixation point preceding them in the same position. The subject then had to decide as quickly and accurately as possible whether or not the presented words were a correct English phrase. The computer recorded each subject´s reaction time for each stimulus phrase from the onset of its presentation. The items were presented randomly in a different order at each test run. After each response, the participant immediately received feedback telling him whether his answer was correct. The feedback also provided information on the exact reaction time in milliseconds and the percentage of correct answers for the experiment at that point.

6.2.5 Apparatus

In order to program the phrasal grammatical judgement experiment, the software package EPrime was used again (see description in section 6.1.6). The test functions in exactly the same way as the lexical decision task described above; the only difference is in the stimulus items. This time the stimuli were not single words but two-word-phrases.

The experiment was run on regular PCs with an Intel Pentium III processor and 1024x768 pixel, 17-inch monitors positioned about 50 cm away from the subjects. Reactions were registered on keyboards that fed to the E-Prime reaction time recording device. The stimuli were presented in a bold 18-point black Courier New font in the centre of the screen. No additional software was needed; the E-Prime software served as the core of the phrasal grammatical judgement task.

6.2.6 Test design

Like the LD task, the phrasal grammatical judgement task was an on-line task. The task of judging correctness was merely an excuse to elicit real-time recognition effects for factors hidden in the design of the task. Participants were informed that the purpose of the experiment was to decide quickly and accurately between correct and incorrect phrases. This statement obscured the researcher´s genuine intention in this on-line experiments. Based on the reaction times, I could then compare various factors that are part of the phrase judgement process and not subject to the participants´ control. I can, for instance, compare the reaction time between conceptual dual nouns and conceptual plural nouns with the same determiner. In the subsections below I describe the considerations involved in selecting the stimulus material in order to fit the rationale (see section 6.2.1) of the task.

6.2.6.1 Mode of linguistic input

In section 6.1.7.1 I briefly discussed the mode of the stimulus input. I will not repeat the arguments here, since they are very similar for this experiment.

Furthermore, the phrasal grammatical judgement task is not a decision task. The point at which the subject recognises the word is not the issue in this task. Subjects need to determine whether or not a given phrase is grammatically correct. In this case, it is of no concern whether the phrase is presented visually or auditorily. It is, for instance, impossible to determine that the phrase “he speak” is incorrect before the complete phrase has been processed. Likewise, it is impossible to determine that the phrase “two goats” is correct before the complete phrase has been processed. Hence, the task is a different one and requires different attention processes. The issue of the mode of input is therefore less important.

6.2.6.2 Selecting the stimulus material

The stimulus words for this experiment had to be carefully selected for their conceptual *numerosity* (cf. section 6.2.1). For the conceptual duals, I thought of all possible entities that occur in pairs. Many of these are body parts like “legs”, “hands”, “eyes”, and “ears”. I made a list that also included the nouns that take a grammatical dual in Hebrew, as listed in a concise and informative overview by Tobin (2000). These include: “pants”, “shoes”, “places”, “days”, etc. Having compiled a long list of items, I dropped those that do not have a dual connotation in English, like: “places” or “days”. Left with a list of about 50 items, I checked the log-frequency (see section 6.1.7.2.1) of both the singular and plural forms of these nouns. As in the previous experiment, I obtained the frequency information from the CELEX database (Baayen, Piepenbrock & Gulikers, 1995). The frequencies of both singulars and plurals were fused into one file with the help of a special piece of software that facilitated the extraction of both frequencies at the same time.

The frequency information on both forms was necessary in order to match the stimulus items for different dominance levels. If I had not considered *dominance*, an eventual RT effect could erroneously be diagnosed as a numerosity effect, rather than an artefact caused by dominance. Dominance is a formal property whereas numerosity is caused by mental association. Therefore, I assured that half of the dual concept target items were plural-dominant, and half were singular-dominant. The target items for the conceptual duals are:

- singular dominant dual: *hands, cheeks, wings, breasts, ankles, elbows*

- plural dominant dual: *eyes, legs, shoes, lips, socks, boots, lungs*

I selected the plural-concept nouns by searching the list of plural-dominant nouns for words that had a clear plural connotation in the sense of referring to more than two and preferably to a large quantity. Nouns included in this list were, for instance “teeth”, “ribs” and “years”. Eventually, I had an equal number of conceptual plurals and conceptual duals. The conceptual plurals, by definition, were plural dominant. The target items of coneptual plurals are:

tears, teeth, miles, hours, trees, toes, fingers, bees, nuts, rays, ribs, herbs

I then used both the conceptual-dual nouns and the conceptual-plural nouns to build noun phrases that had both a dual and a plural determiner, as I described in section 6.2.1. I distributed the target items evenly between two test versions in order to avoid listing the same noun twice with two different determiners. In the end, there were 25 target items (13 conceptual duals and 12 conceputal plurals). Each test version contained all 25 of these target items but with different determiners. I also included some correct noun-phrase and verb-phrase fillers, as well as an equal number of incorrect noun- and verb-phrase distractors. A complete list of all stimuli is included in the appendix. The following 195 stimulus items were included in the phrasal grammatical judgement task; the bold-face examples are the target items.

- **13 dual-concept noun phrases with numeral determiner, correct** (e.g. *two shoes*)
- **12 plural-concept noun phrases with numeral determiner, correct** (e.g. two herbs)
- 13 singular-concept noun phrases with numeral determiner, correct (e.g. *two loves*)
- 38 noun phrases with numeral determiner, incorrect (e.g. *seven oak*)
- 38 verb-phrase fillers, correct (e.g. *he stirs*)
- 38 verb phrases, incorrect (e.g. *she guess*)
- 13 noun phrases with pronoun and noun, incorrect (e.g. *she country*)

- 15 noun phrases with non-numeral determiner, correct (e.g. *a river*)

- 15 verb phrases with determiner, incorrect (e.g. *the grows*)

= a total of 195 items

The two alternate test versions both included all 195 stimulus items. The two versions differed solely in the numeral determiner of the 25 target items. These items were matched with either a dual-numeral or a plural-numeral determiner. The two noun-phrase matches were distributed pseudorandomly (random distribution generated by a computer computation) among the two versions in order to ensure an equal number of dual and plural determiners in both experiments.

7 *Experimental findings*

The previous chapter explained the purpose of the the two experiments, it gave an overview of the most prominent findings observed in previous experiments similar to my own and provided some insight into the general setup of the two tasks used in this study. In particular, the criteria for stimulus selection as well as the procedure of the task itself were explained. This chapter presents the raw results of both experiments categorised for the two main groups of native and non native speakers and according to the variables introduced in Chapter 6.

7.1 Experiment I: Lexical decision

7.1.1 Procedure

The analysis of the subjects´ reaction times was carried out using two pieces of software: EPrime Data Aid and SPSS 11.0 statistical software. In order to filter out all the unusually high or low outliers, I calculated the mean reaction time and the standard deviation (SD) of each subject. All RTs more than two SDs above or below the mean for each subject were filtered out in order to avoid unusual RTs that might have been caused by distraction or guessing. The RTs filtered out were analysed as missing values in the statistical analysis. Of the 208 subjects tested, 22 had to be excluded from the analysis because they either talked during the experiment or became distracted or because they failed to finish. The analysis therefore includes 186 subjects: 55 native speakers and 131 German learners of English.

A 2x3x2 mixed randomized-repeated ANOVA was performed on the subjects´ reaction times. Each stimulus word had a value for each of the three variables: number (singular or dominant), frequency (high, mid and low) and dominance (singular dominant or plural dominant). The fourth variable, "relativity" (cf. 6.1.7.2.3), was nested in the dominance variable and was not treated as an independent variable. The multivaried repeated ANOVA compares all of these variables and their interaction within all subject groups but also between groups, for instance native speakers versus learners. For the purpose of finding processing differences between natives and learners, it was important to compare the two large subject pools (native versus nonnative). When comparing the native speakers with the learners, the native English A-level pupils and the soldiers were taken together as one native speaker group. In case of any differences in these groups, individual outcomes for the separate native speaker groups are mentioned. The comparison between the different classes (grades 7 through 12) in the German gymnasium of the ANOVA made it possible to look at each group individually.

7.1.2 Native speakers

7.1.2.1 number: singulars versus plurals

For the 55 native speakers, sixth form pupils and soldiers taken together, the overall mean reaction time was 613 ms (SD 12), an average of 645 ms for the native-speaking soldiers, and an average of 573 ms for the native-speaking A-level pupils in the sixth form at the British army school in Gütersloh. The mean RT difference between the two native-speaker groups provided a large contrast that is statistically significant F (1,53) = 11.86, $p < .001$, eta2 = .183 and will be discussed together with all other results in a later section.

There was no overall number effect (a significant difference between RTs to singulars and plurals) for the natives as a uniform group. That is, on the whole, they did not process singulars faster than plurals. Considering the groups separately, it is striking that the reaction times of the sixth formers show no number effect at all, whereas the soldiers´ RTs come close to showing a number effect although it fails to reach significance (F (1,30) =

3.99, $p < .055$, eta2p = .118 at a confidence interval of 95%). The soldier´s response time latencies for the singular nouns have a mean RT of 636 ms (SD 16). The singulars are almost significantly faster than the plural nouns. The response time latencies for the plural nouns have a mean RT of 653 ms (SD 15). For the soliders, the number effect is not significant but it is worth considering.

7.1.2.2 Frequency: High-, mid- and low-frequency nouns

The native speakers generally responded to the nouns of different frequency values with significantly different reaction times. The statistical analysis of the RTs for the frequency variable showed highly significant frequency effects for the native speakers, both soldiers and students. This indicates that they processed the three different frequency levels of nouns with significantly different mean time latencies.

The reaction times for the different frequency levels of the sixth formers are convincingly different with a step-wise increase of reaction time per level. The mean RT for high-frequency nouns is 550 ms. The mean RT for mid-frequency nouns is 570 ms and the mean RT for low frequency nouns is 600. The RTs for the frequency levels are significantly different from each other: F (1,46) = 13.33, $p < .001$, eta2p = .37. The biggest difference of 50 ms can be found between high and low frequency nouns. The sixth formers´ response latencies were fastest in the category of high frequency nouns and seemed to increase regularly in RT for mid and low frequency nouns.

For the native-speaker soldiers, this regular increase cannot be observed. The soldiers responded to high-frequency nouns with a mean RT of 638 ms and to mid-frequency nouns with a mean RT of 626 ms. Although, on average, the RTs for the mid-frequency nouns are faster than for the high-frequency nouns, this difference is negligible and not significant. The low-frequency nouns, on average, showed the longest reaction time latencies with a mean RT of 670 ms. The mean RT measure for the group of low-frequency nouns is significantly slower than for the high- and mid-frequency groups.

Table 6. Mean effect for the frequency factor within the group of British soldiers

Pairwise Comparisons of three different frequency levels [b]

Measure: MEASURE_1

(I) FREQ	(J) FREQ	Mean Difference (I-J)	Std. Error	Sig. [a]
high	mid	12.716	12.298	.309
	low	-31.130*	15.174	.049
mid	high	-12.716	12.298	.309
	low	-43.846*	10.741	.000
low	high	31.130*	15.174	.049
	mid	43.846*	10.741	.000

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

b. group = British soldiers

For native speakers, this lack of frequency effect between the high- and mid- frequency levels may not be too surprising since mid-frequency words in a native language are relatively familiar and sufficiently frequent to the native speakers, whereas they might just not be familiar enough to non-native speakers. Low-frequency nouns may be less psychologically salient in the real native language environment and might therefore be significantly different from the other two frequency groups. For the British soldiers, it seems, low-frequency nouns were sufficiently infrequent as to influence processing quality when compared to high- and mid- frequency nouns.

7.1.2.3 Interaction number x dominance

The interaction between the number and the dominance variable combines the effect of the processing difference between singular and plural nouns (number variable) and the frequency-based dominance variable. The dominance variable is determined by the relative number of occurrences of the noun in either the singular or the plural form. If a noun occurs more often in the plural than in the singular, the noun is plural dominant. The dominance effect has been experimentally demonstrated in earlier experiments (cf 5.4) and will also play a role in the results of this LD task.

7.1.2.4 interaction with absolute dominance

The quality of the dominance variable alone does not influence processing quality. Whether a word appears more often in the singular or in the plural form is irrelevant if one wants to compare the reaction times of two singular forms only. It is only in the interaction with the number variable that dominance can be assessed as a psychologically valid factor. If we are to make sense of the frequency of occurrence of the singular form in relation to the frequency of occurrence of the plural form, we must compare the reaction times of both the singular and plural forms in both conditions. We can expect a singular-dominant noun to be processed faster in its singular form than in its plural form, as previous experiments have shown (cf. 5.4). A plural-dominant noun, however, can be expected to be processed faster in its plural form than in its singular form.

The sixth formers showed a highly significant (F (1,23) = 4.9, p = .037, eta2p = .18) effect for the interaction between number and dominance. Interestingly, this effect showed very strongly for the plural-dominant condition (Figure 2: the green line is steeper than the blue line); on average the pupils processed plurals almost 20 ms faster than singulars. Figure 2 illustrates the RT differences between the singular dominants on the one hand and the plural dominants on the other hand.

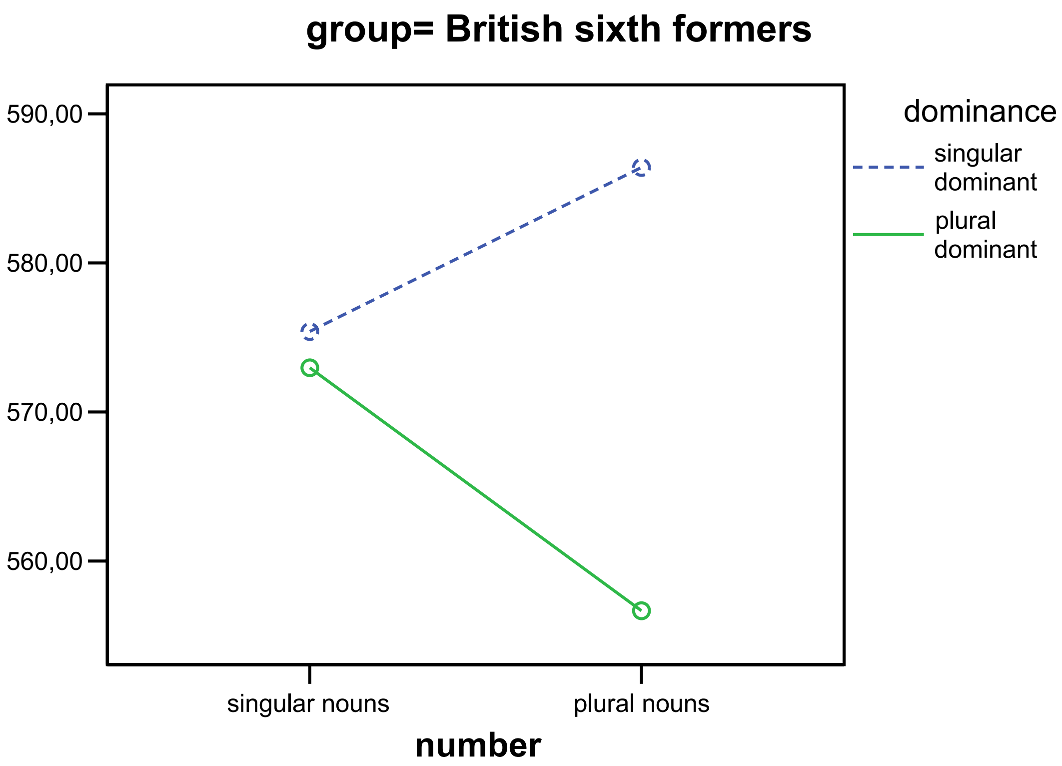
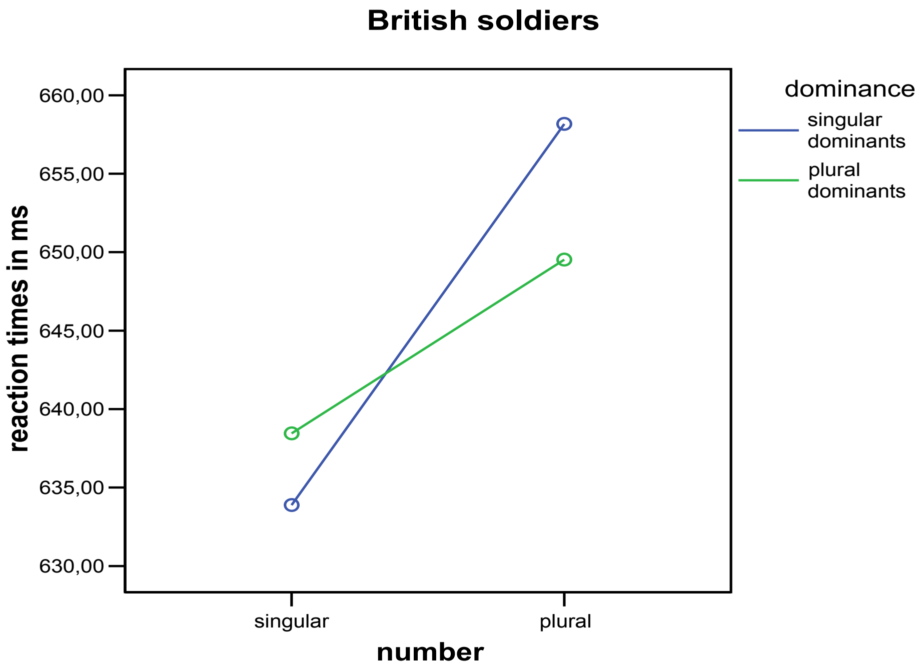


Figure 2. *Interaction number x dominance for British sixth formers*

The most interesting finding here is the fact that dominance as a factor seems to play a role only for plural nouns. For singular nouns it is of no interest whether the word is singular or plural dominant. The lines only diverge for plural nouns, as can be seen on the right-hand side of Figure 2. For singular nouns the RTs are almost identical regardless of the dominance value. The dominance value seems to enhance the potential of the processing latencies of plurals, either slowing down or accelerating the process of comprehending plural nouns (on the right-hand side) with respect to the singular forms.

For the British soldiers, the interaction between number and dominance was not significant (see Figure 3). Within the plural dominant condition, there seemed to be a reverse effect; on average they processed singulars about 10 ms faster than plurals.

Figure 3. *Mean effect for the interaction number x dominance for British soldiers*



The graph shows that the interaction between number and dominance is not significant. If it were significant, the green plural dominant line would most likely be going down, representing faster reaction times for the plural nouns than for the singular nouns. There is a relatively large interaction effect for the singular dominant condition (the red line is much steeper here compared with that for the sixth formers).

It is important to note that both native speaker groups show a larger effect for the plural nouns than for the singular nouns. The plural nouns seem to be more likely to be affected by the dominance variable. The reaction times for singular nouns do not display any important difference. Dominance seems to be decisive for plural noun forms, not for singulars.

Although it is hard to draw a general conclusion from the data of both the native speaker groups for the interaction between number and dominance, the results for the singular dominant condition are clear. In both native speaker groups, the singular nouns in the singular dominant condition were processed about 10-20 ms faster than the plural nouns. This is a stable effect for the native speakers. The relevance of this finding will be discussed in a later section.

7.1.2.5 Interaction with relative dominance

As pointed out in section 6.1.7.2.3, the dominance variable was split into two subvariable groups, those with relatively high dominance and relatively low dominance. Since dominance per se may not be decisive for processing quality in cases where a word is only marginally plural dominant, one has to consider the degree of dominance. As explained in detail in section 6.1.7.2.3, only singular and plural dominants whose dominant number (.e.g. singular) exceeded the defined log-frequency range of the respective opposite number (e.g. plural) are considered of relatively high dominance. Using this new definition of dominance, I will now present the statistics for the singular- and plural-dominant nouns with relatively high dominance only.

The results of the ANOVA to determine the interaction number x relatively high dominance for the sixth formers are significant ($F(1, 27) = 4.2$, $p < .050$). Compared to the interaction with absolute dominance values (Figure 2), this effect has not changed. However, the effect has become much stronger for the plural nouns. Hence, the overlap between the two endpoints of the lines on the right-hand side of the graph is larger than in Figure 2. As observed in the first measurement, the interaction effect is not very large for the singular nouns. That is, there is hardly any difference between the singular nouns of either dominance condition.

Figure 4. Interaction number x dominance (relatively high dominance only) for the sixth formers

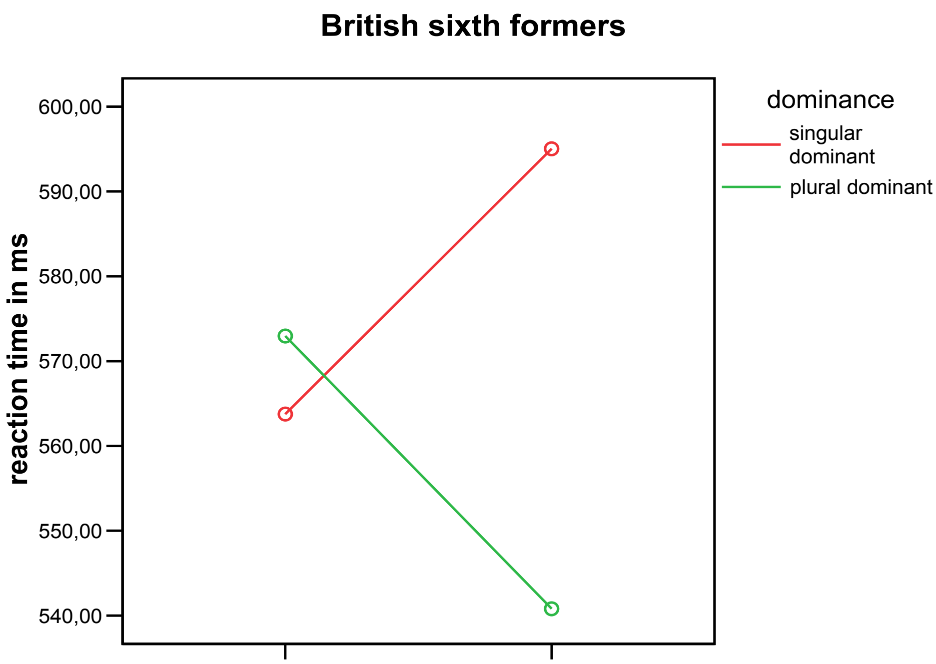
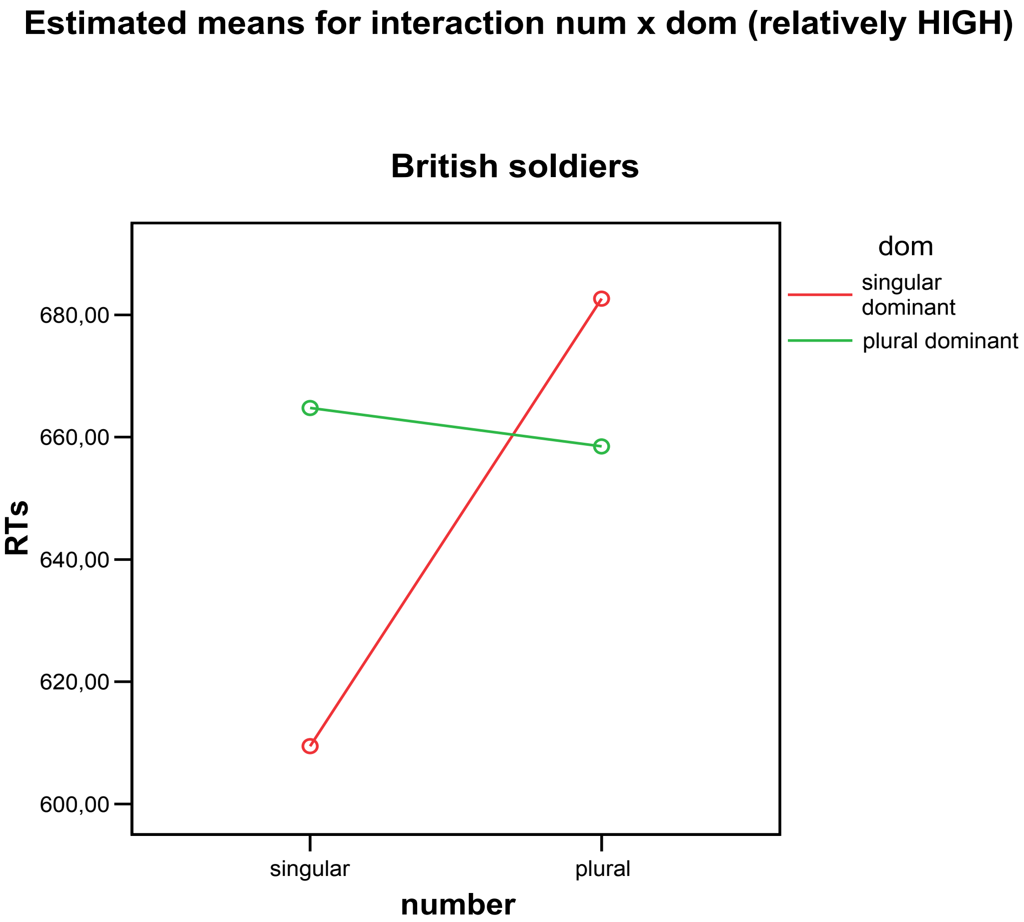


Figure 5. *Interaction number x dominance (relatively high dominance only) for the British soldiers*



According to the ANOVA, the performance of the British soldiers is similar to that of the pupils (see Figure 5). The interaction effect between the high-dominance variable and the number variable is significant, as for the British pupils F (1, 33) = 4.3, p < .045. This interaction effect increased noticeably when I removed the relatively low dominant values from the analysis. This increase was most noticeable for the plural dominants. The reaction times to plural-dominant nouns remain more or less the same regardless of whether they were in the singular or plural form. The singular-dominant nouns, however, elicited an enormous range of responses, from roughly 610 ms to 680 ms. At the same time, for both singular and plural noun processing, it matters whether the noun in question is singular- or plural-dominant with relatively high dominance. Unlike the pupils, the soldiers also showed this effect for the singular nouns. In fact, it is the singular-dominant nouns that act as a catalyst to either speed up or slow down the reaction times of singulars and plurals.

7.1.3 Learners of English

7.1.3.1 number: singulars vs plurals

The learners of English, in class 7 through 12 showed different behaviour patterns with respect to their reaction times for singular and plural nouns. In most of the groups (classes 12, 10, 9 and 8), the reaction times for singular and plural nouns did not differ significantly. In this respect, these learners of English did not differ substantially from the native speakers.

Within the learner group, however, two classes did have significantly faster mean reaction times for singular nouns than for plural nouns, which means they displayed a number effect. These two groups are class 7 and class 11. For class 7, there is a difference of almost 20 ms between the mean RT for singulars and the mean RT for plurals. This difference is highly significant $F\ (1, 24) = 4.9\ p < .036$. In total, these students´ reaction times were rather slow when compared to the natives. The reaction times of the second group (class 11) also show a significant number effect $F\ (1, 15) = 5.4, p < .034$. The mean RT for the singular nouns is 617 ms and the mean RT for plural nouns is 645 ms. Overall, these reaction times are much faster than the RTs of the younger learners of class 7, both the singulars and the plurals.

The most striking observation here is the significant effect for class 7 and class 11, and the lack of any such effect for all the other learner groups. The number effect therefore cannot be strictly assigned to either the beginning learners or the more advanced learners because of the large difference (4 years) between class 7 and class 11. No significant number effect could be found in any of the other classes even though a difference between the mean RT for singulars and the mean RT for plurals is apparent in all learner groups. However, these are simply mean reaction time values and, statistically, these differences do not seem to play a role. At first glance one would diagnose a number effect for all 6 classes but the statistics prove this effect to be genuine only for class 7 and class 11. This seemingly arbitrary effect does raise a question about the role of the number effect for either native or non-native processing.

7.1.3.2 Frequency: High-, mid- and low-frequency nouns

The frequency effect is an easy one to detect. The analysis of the reaction times of the learner groups reveals a highly significant frequency effect for all of them, as for the native speaker groups (cf. 7.1.2.2). This means that reaction times for high-, mid- and low-frequency nouns are significantly different from each other. This effect is a very reliable and a very stable one. The only observation that is interesting in this context is the nature of this frequency effect. That is, there may be significant differences between high- and low-frequency nouns, between high- and mid-frequency nouns, and possibly between all frequency levels. Because the statistics for the sublevels may differ from group to group, I now look at them in close detail.

Class 12, the highest class with the oldest learners shows a highly significant overall frequency effect: F (2, 30) = 21.48, p < .000, eta2p = .589. Table 7 shows the statistics for the different frequency levels. It is obvious that all the differences between the frequency levels are highly significant.

Table 7. *Frequency effect for class 12*

Pairwise Comparisons[b]

Measure: MEASURE_1

(I) FREQ	(J) FREQ	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
high	mid	-49.655*	10.604	.000	-72.257	-27.054
	low	-114.184*	22.090	.000	-161.268	-67.101
mid	high	49.655*	10.604	.000	27.054	72.257
	low	-64.529*	17.757	.002	-102.377	-26.681
low	high	114.184*	22.090	.000	67.101	161.268
	mid	64.529*	17.757	.002	26.681	102.377

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equiva adjustments).

b. GRUPPE = Klasse 12

Table 8. Frequency effect for class 11

Pairwise Comparisons[b]

Measure: MEASURE_1

(I) FREQ	(J) FREQ	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	Upper Bound
high	mid	-41.279*	12.458	.005	-67.833	-14.725
	low	-63.915*	18.476	.004	-103.296	-24.535
mid	high	41.279*	12.458	.005	14.725	67.833
	low	-22.636	19.479	.263	-64.155	18.883
low	high	63.915*	18.476	.004	24.535	103.296
	mid	22.636	19.479	.263	-18.883	64.155

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equival adjustments).

b. GRUPPE = 11. Klasse

For the learners in class 11 the statistical analysis of the reaction times reveals a highly significant overall frequency effect: F (2, 30) = 7.2, $p < .003$, eta2p = .324. For this learner group, Table 8 shows that this significant difference is not distributed evenly among the three frequency levels. Whereas the RTs for both high- and mid-frequent nouns, as well as those for high- and low-frequency nouns differ significantly from each other, the RTs for mid- and low-frequency nouns are not significantly different. This means that for class 11, only high-frequency nouns are significantly frequent to evoke a qualitative processing difference from the RTs of the other two frequency levels.

The analysis for class 10 also shows a highly significant overall frequency effect F (2, 38) = 20.99, $p > .000$, eta2p = .525. This is the case for all frequency levels as shown in Table 9. Again, the RTs between all three levels differ significantly, and there seems to be a gradual digression in mean reaction times for the three groups, as was true for class 12.

Table 9. Frequency effect for class 10

Measure: MEASURE_1

(I) FREQ	(J) FREQ	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interva Difference[a]	
					Lower Bound	Upper Bo
high	mid	-42.962*	10.682	.001	-65.318	-20.6
	low	-94.330*	17.134	.000	-130.191	-58.4
mid	high	42.962*	10.682	.001	20.605	65.3
	low	-51.368*	15.163	.003	-83.104	-19.6
low	high	94.330*	17.134	.000	58.469	130.1
	mid	51.368*	15.163	.003	19.632	83.1

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (eq adjustments).

b. GRUPPE = 10. Klasse

Again, the analysis of the reaction times of the 8th and the 9th classes show a highly significant overall effect: $F\ (2, 40) = 29.03, p < .000$, eta2p = .592 and $F\ (2, 64) = 23.64, p < .000$, eta2p = .425 respctively. It also shows significant differences between all three frequency levels. Tables 10 and 11 show the statistics on the different levels for the two learner groups.

Table 10. Frequency effects for class 9

Pairwise Comparisons[b]

Measure: MEASURE_1

(I) FREQ	(J) FREQ	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	Upper Bound
high	mid	-37.877*	6.061	.000	-50.223	-25.530
	low	-56.606*	9.566	.000	-76.091	-37.121
mid	high	37.877*	6.061	.000	25.530	50.223
	low	-18.729*	9.102	.048	-37.270	-.189
low	high	56.606*	9.566	.000	37.121	76.091
	mid	18.729*	9.102	.048	.189	37.270

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equival adjustments).

b. GRUPPE = 9. Klasse

Table 11. Frequency effects for class 8

Pairwise Comparisons[b]

Measure: MEASURE_1

(I) FREQ	(J) FREQ	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval Difference[a] Lower Bound	Upper Bou
high	mid	-32.877*	10.573	.006	-54.932	-10.82
	low	-116.752*	19.037	.000	-156.464	-77.04
mid	high	32.877*	10.573	.006	10.823	54.93
	low	-83.875*	16.582	.000	-118.465	-49.28
low	high	116.752*	19.037	.000	77.041	156.46
	mid	83.875*	16.582	.000	49.285	118.46

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equi adjustments).

b. GRUPPE = 8. Klasse

The analysis of the RTs of the youngest group, class 7, also reveals a significant overall frequency effect F (2, 48) = 14.62, $p < .000$, eta2p = .379. Compared to all the other groups, the reaction times of class 7 are highest for all frequency levels. As Table 12 shows, the mean reaction times for the mid- and low-frequency groups are similar; therefore only the mean RTs for the high-frequency words differ significantly from the other two groups, as shown in Table 13.

Table 12. *Reaction times for frequency levels in class 7*

Estimates[a]

Measure: MEASURE_1

FRE	Mean	td. Erro	% Confidence Inter	
			ower Bour	pper Bour
high	49.902	17.741	613.286	686.518
mid	15.829	25.046	664.138	767.521
low	13.564	23.910	664.216	762.913

a. GRUPPE = 7. Klasse

Table 13. *Frequency effects for class 7*

Pairwise Comparisons[b]

Measure: MEASURE_1

(I) FREC	(J) FREC	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval f Difference[a]	
					Lower Bound	Upper Boun
high	mid	-65.928*	13.918	.000	-94.653	-37.202
	low	-63.663*	14.108	.000	-92.781	-34.545
mid	high	65.928*	13.918	.000	37.202	94.653
	low	2.265	13.496	.868	-25.589	30.119
low	high	63.663*	14.108	.000	34.545	92.781
	mid	-2.265	13.496	.868	-30.119	25.589

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equiv adjustments).

b. GRUPPE = 7. Klasse

As was true for the number effect, it is again the reaction times of the learners in the 7th and the 11th class that are similar. Both groups lack a frequency effect between mid- an low-frequency nouns. In these two groups, the mid-frequency nouns are not prominent enough to evoke a qualitative processing difference from the low-frequency nouns. Only the high-frequency nouns are salient enough for processing and evoke a significant effect.

7.1.3.3 Interaction number x dominance

As I did for the native speakers, I will present my analysis of the interaction between the number and the dominance variable in two steps. First, I will consider the absolute dominance value in the interaction analysis. Second, I will include in the analysis only the highly dominant (high relativity) nouns.

7.1.3.4 Interaction with absolute dominance

The analysis of absolute dominance, by itself, revealed almost no interaction effects in any learner group. The only interaction effect was in the reaction times of the class 9 learners: F (1, 32) = 5.87, $p < .021$, eta2p = .155. Singular dominant nouns were processed much faster in the singular than in the plural (mean RT of 619 ms for singulars and mean RT of 645 ms for plurals). And plural dominant nouns were processed faster in the plural form than in the singular form (mean RT of 611 for plurals and mean RT of 622 for singulars).

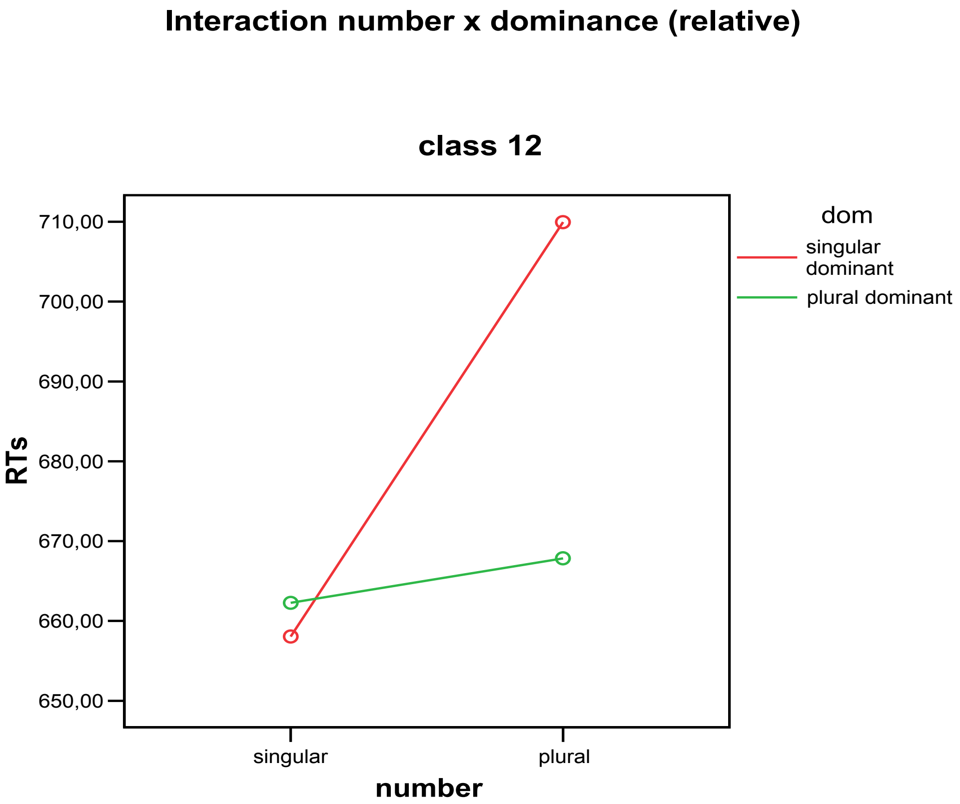
Four classes, (class 7, 8, 10 and 11) however, showed a different pattern for singular- and plural-dominants. Here, singular dominance does influence the processing quality and slows down the reaction times for plural nouns. Plural-dominance does not have any influence on the processing quality; the RTs for singulars and plurals are similar within this category. Since the effect can only be detected for the singular dominants but not for the plural dominants in the above groups, one cannot speak of a true interaction effect. In order to illustrate how such an effect would manifest itself, I will present the mean RTs of class 8 as an example. The mean RT for plural-dominant nouns is 656 ms for singular nouns and 655 ms for plural nouns. This difference of 1 ms is irrelevant. The mean RT of singular-dominant nouns ranges between 677 ms for singular nouns and 698 ms for plural nouns. Hence, students process the singular dominants much more quickly when the noun is presented in its singular form.

7.1.3.5 Interaction with relative dominance

The results of the statistical analysis for the interaction between number and relatively high dominance are very interesting when compared with the native speaker difference between the two analyses. Although we found an interaction effect for one learner group in the previous analysis, none of the learner groups seemed to produce any interaction effect in this second analysis. The relatively high-dominants included in the second analysis are a subgroup of the total number of dominants in the previous analysis (see 7.1.3.4). The samples were probably not large enough to evoke any interaction effect, even in the group that produced one in the earlier analysis.

The only group that came close to producing a significant interaction effect is class 12, the oldest learner group: F (1, 23) = 2.73, $p < .088$, eta2p = .112. Figure 6 shows a clear difference in response behaviour between singular and plural dominant nouns. The singular dominant nouns evoke a very clear RT difference between singular forms and plural forms. The plural dominants hardly do have any influence on the RTs for either singular or plural forms. Figure 6 illustrates how large this effect seems to be for singular dominants. The red line is very steep and indicates a large difference between RTs for singular and plural forms. The red line is stable throughout all learner groups. However, the green line (plural dominants) in the data for class 12 shows hardly any effect and it varies from one learner group to another. We can therefore conclude that for the learners there is a common effect for singular dominants only and not for plural dominants.

Figure 6. Interaction number x dominance (relative) for class 12



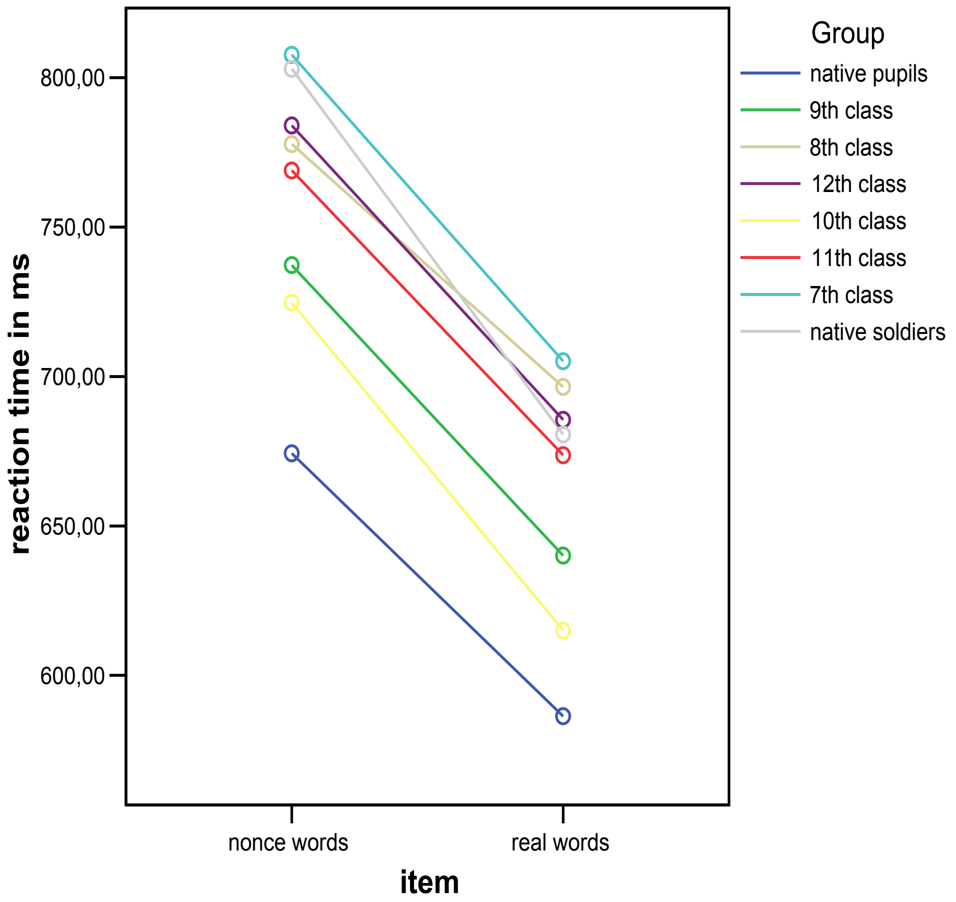
7.1.4 Overall reaction times

Figure 7 shows the general RTs of all groups. Since the real and nonce words altogether make up 100% of the stimulus items of the test, one can take these times as an indicator of the general reaction time for each group. The blue line at the bottom of Figure 7 is the line for the British native sixth formers. Clearly, they responded most quickly to the test items. Surprisingly, the soldiers, the other native speaker group, was not among the fastest. They are at the slowest end of the range, together with the German pupils in 7th and 8th class. I did not find a regular decrease in reaction time corresponding to the age of the learner groups. Rather, the two youngest groups, with the least exposure to English language teaching, were also the slowest to respond. However, the oldest learner group (12th class), with the longest exposure to English language teaching, was also one of the slower groups.

The lexical decision task includes real words (among them the target words) as well as nonce words. Although the analysis is not relevant for the claims of this thesis, I am presenting the overall reaction times for real and nonce words for all subject groups. I described this issue in section 6.1.2.2.

Figure 7 immediately reveals that overall, nonce words take longer to process than real words. All groups show this difference of roughly 100ms between nonce and real words in the mean RT for all the 400 stimulus items on each test. In all groups this difference is highly significant: F (1, 196) = 465.1, p < .000, eta2p = .704 This finding seems to be robust and therefore supports the results of previous studies (cf. 6.1.2.2).

Figure 7. Overall reaction times for real and nonce words



7.2 Experiment II: phrasal grammatical judgement

7.2.1 Procedure

The subjects´ reaction times were analyzed using EPrime Data Aid, which then exported the file into the SPSS 11.0 statistical software. As in the first experiment, all unusually high or low outliers were filtered out by calculating the mean reaction time and the standard deviation (SD) for each subject and deleting all RTs that were more than two SDs above or below the mean for that subject. The deleted measurements were analysed as missing values in the statistical analysis.

Of the 169 subjects tested, 9 native speakers and 9 learners had to be excluded from the analysis because they left the experiment before finishing or showed a distinct lack of concentration during test-taking. This included talking to neighbours and looking out of the window instead of at the screen. The analysis therefore includes only 151 of the above 169 participants: 54 native speakers and 97 learners of English.

A 2x2 mixed randomized-repeated ANOVA was performed on the subjects´ reaction times. Each target stimulus noun phrase had a value for its numeral determiner (either dual or plural) and its noun concept (either dual or plural). This resulted in four different noun-phrase conditions, as explained in section 6.2.6.2. The multivaried repeated ANOVA compares these four target conditions and their interaction within all subject groups but also between groups, for instance native speakers versus learners. The data was then split into two parts, in order to obtain all the p-values needed to compare the two conditions within the dual-noun group and all those needed to compare the two conditions within the plural-noun group.

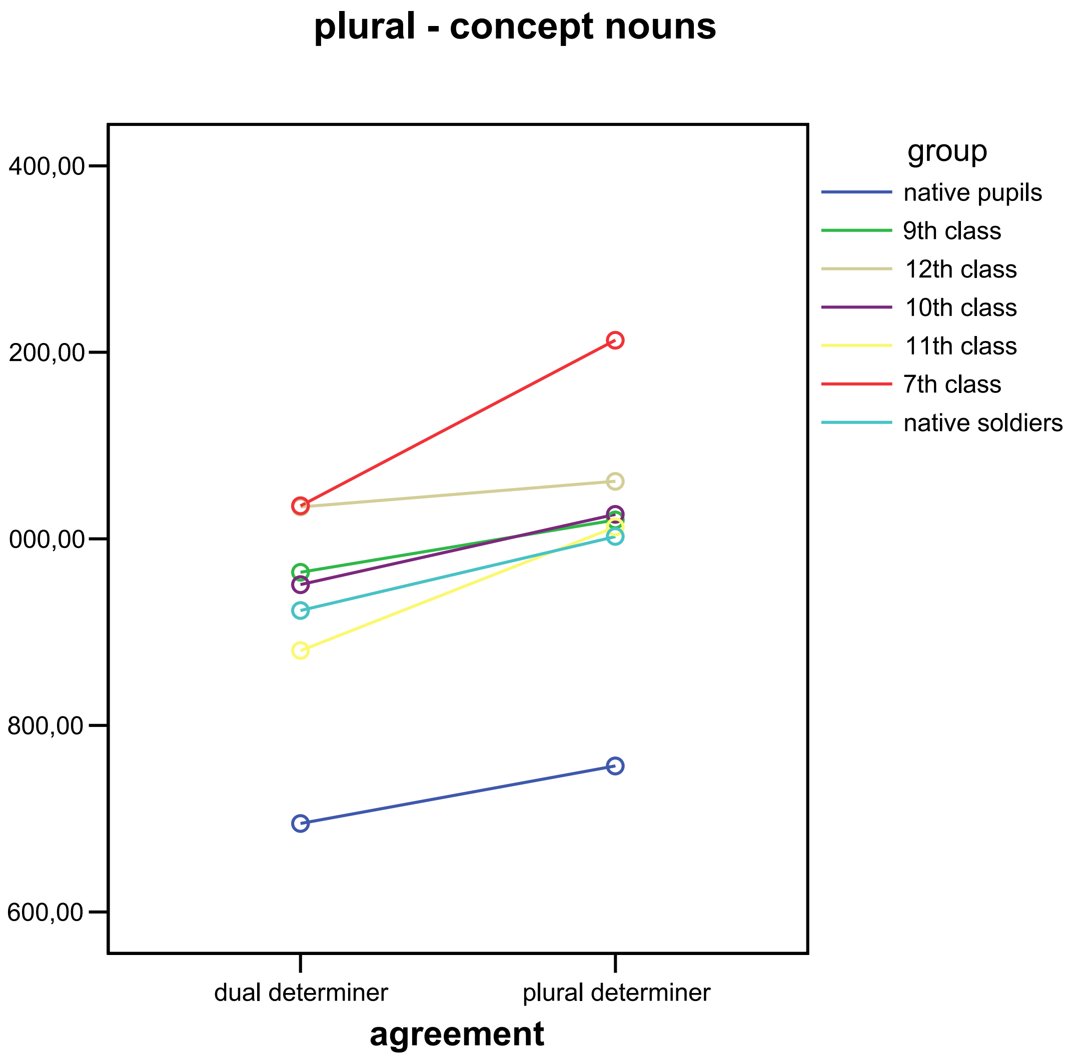
7.2.2 Plural-concept noun phrases

In the phrasal grammatical judgement task, the conceptual plural nouns were matched with either a dual numeral or a plural numeral (see section 6.2.6.2). These two conditions were compared for each group and are presented in Figure 8. As can be seen clearly, the groups all behaved in a similar way. Reaction times for plural-concept nouns with a dual numeral (e.g. *two teeth*) generally seem to be slightly faster than reaction times for plural-concept nouns with a plural numeral (e.g. *ten teeth*). These differences are highly significant only in the case of class 11, indicated by the yellow line in Figure 8: $F(1, 14) = 16.4$, $p < .001$, eta2p = .539. The reaction times of the British pupils come close to showing a significant agreement effect, with $F(1, 25) = 4.1$, $p < .054$, eta2p = .141. For all the other groups, the difference between the two agreement conditions did not seem to be significant. A tendency towards an agreement effect can be detected in the reaction times for class 10: $F(1, 19) = 3.64$, $p < .072$, eta2p = .161, although the difference between the two mean RTs in this case might also be purely accidental.

Looking at the lines in Figure 8, one could assume that because of the statistically significant difference between the two mean RT values for the 11th class (yellow line), the mean RTs for the 7th class (red line) should also be significantly different. This, however, is not confirmed by the statistical analysis. Instead, the analysis of the data of the 7th class reveals that the standard deviation (SD) for the two mean values is between 500 and 600 ms. This SD value is very high, suggesting that the subjects were inconsistent in their responses. Thus the values for this class are unreliable and the findings cannot be called significant.

Summing up, I found that for one learner group the data reveals a significant difference between the two mean reaction times and for one native-speaker group it came close to showing a significant effect with a processing preference for the dual determiner. The data of the other four learner groups and of the British soldiers do not show a significant agreement effect in the responses to the plural-concept nouns.

Figure 8. *RTs for the different agreement conditions for the plural concept nouns*



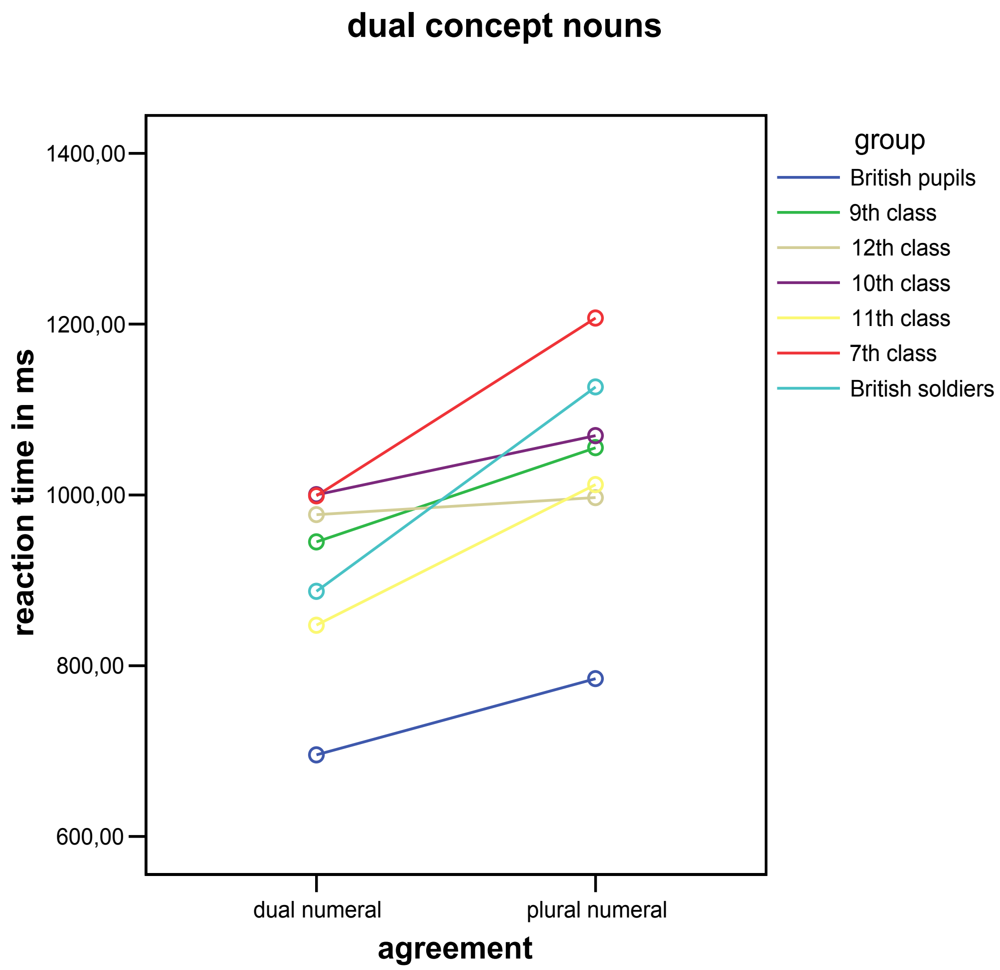
7.2.3 Dual-concept noun phrases

Like the concepual-plural nouns, the conceptual-dual nouns were matched with dual and plural numerals (cf. section 6.2.6.2). Figure 9 shows the mean RTs for the dual-numeral / dual-concept and the plural-numeral / dual-concept conditions for each group.

At first sight, the lines in Figure 9 look similar to those in Figure 8 for the plural-concept nouns. On the whole, the mean reaction times for the dual-numeral / dual-concept nouns seem faster than those for the plural-numeral / dual-concept nouns. Only class 12 showed different mean response times (grey line). The statistical analysis of the data for class 12 (grey line) reveals that, indeed, there is no statistical difference between the two agreement conditions. Surprisingly, this is also the case for the 10th class (purple line): there is no statistically significant difference in their mean response times between the two agreement conditions. The analysis of the mean RTs for the 9th class (green line) yields an agreement effect that comes close to being significant: F (1, 17) = 4.3, $p < .054$, eta2p = .201. The mean RTs of all the other groups do yield agreement effects that are clearly significant. The native pupils showed a highly significant difference in the two dual-concept conditions: F (1, 25) = 15.7, $p < .001$, eta2p = .386. The native soldiers also showed a highly significant effect: F (1, 24) = 18.9, $p < .000$, eta2p = .441. And the remaining two learner groups, namely class 11: F (1, 14) = 8.84, $p < .010$, eta2p = .387, class 7: F (1, 25) = 4.96, $p < .035$, eta2p = .166 also clearly showed a significant agreement effect.

Summing up, for two learner groups the data show a non-significant agreement effect. For one learner group the effect is very close to being significant. The mean RTs of the remaining two learner groups, and of the native speaker groups, do show a highly significant agreement effect for the dual-concept nouns with a processing preference for dual determiners.

Figure 9. *RTs for the different agreement conditions for the dual-concept nouns*



8 Discussion

8.1 Summing up

This thesis set out to explore the relationship between lexicon and grammar in the inflectional processing of pluralised nouns in English as a native and as a second language. As discussed in the review of the literature in the first chapters, grammar and lexicon have widely been accepted as two distinct processes, both in linguistic processing time and processing location. The literature provides ample empirical evidence to suggest such a distinct functional and partly also locational difference between the two. This difference is reflected in the traditional distinction between inflectional and derivational morphology (Perlmutter, 1988; Badecker & Caramazza, 1989; Schriefers, Friederici, and Graetz, 1992), in the distinction between inherent and contextual morphemes (Booij, 1993; Pienemann, 1998; Tobin, 2000; Hirtle, 2000; Hay, 2001) or in the distinction between regular and irregular inflectional morphemes (Sonnenstuhl, Eisenbeiss & Clahsen, 1999; Lück, Hahne & Clahsen, 2001; Ullman, 2004) which shows that there is room for different processing behaviour within the realm of inflectional morphology.

The same is true for typological differences: inflections may not behave the same in different languages. The examples presented in chapter 5 focus mainly on Finnish and Italian (cf. 5.2.2) and show that processing depends largely on issues such as frequency or productivity rather than on their category as inflections or derivations. Recently, a functional distinction between the two types has also been confirmed in neuroimaging studies. Ullman (2004) claims that there are two brain memory systems at work which are functionally separated and largely correspond to a distinction between rote learning and rule application. For English pluralisation, this traditional distinction between lexicon and grammar in processing suggests that a noun in its singular form is accessed in the mental lexicon. Only then, is an extra-lexical rule applied to add an inflectional morpheme (cf. 2.1). This view sees pluralisation as a strictly grammatical event, based on post-lexical rule application.

This thesis, however, shows that there may be a certain degree of overlap between lexicon and grammar in particular processing situations. Very much like the discussion about the role of the first and the second language lexicon and its degree of modularity in bilinguals by Weinreich (1953) and continued by various scholars such as de Groot (1993) and Kroll (1993) (cf. 2.3.2 ff) I suggest that to determine the correct account for lexical and grammatical processing one must allow for idiosyncratic differences and developmental stages within an individual speaker. The modules are believed to be fuzzy and not clearly demarcated. The work of Sereno and Jongman (1997) and of Hay (2001) casts doubt on the strict dual-route model of morphological processing. Also Ullman (2004) acknowledges that the two brain memory systems corresponding to grammatical and lexical processing are not language specific and not clearly separable in terms of domain specificity (cf. 2.3). And Jones (2004) completely denies that language is unique; instead she advocates for a memory system responsible for semantic processing as opposed to a memory system for episodic memory, regardless of whether it is linguistic. The two linguistic processes therefore need to be regarded as organic and functionally different but allowing different degrees of overlap. Consequently, English pluralisation may not merely be a grammatical process but, rather, a multi-faceted processing event, allowing for various factors to influence the process and render it either more “lexical” or more “grammatical”. The results of the online experiments presented in the previous chapters suggest no distinct demarcation between lexical and grammatical processing, unlike much of the literature presented in the first chapters.

Therefore, the first objective of this work is to define the status of plural processing as either purely grammatical or not strictly so, and second, to find the factors that render plural processing to be either more grammatical or more lexical. The results of the experiments suggest that processing speed depends on factors such as frequency, dominance and conceptual association, which I discuss below.

Since plural processing and the factors influencing plural processing are not restricted to native speakers, we must also consider the same issues for language learners. Language acquisition is a central aspect of this thesis because learners need to develop some degree of separation or overlap between the lexicon and the grammar. I was interested to see at what point in development the processing of inflectional plurals may become native-like. I was also interested to see whether the difference between native processing and non-native processing of plural inflections is qualitatively different. Theoretically, it could be possible to find only quantitative changes between the different levels of learners up to the most advanced ones and a qualitatively different behaviour in the native speakers. There could also be a gradual processing difference between beginning learners, and more advanced learners, up to the native speakers. And if that happens, is such a gradual difference reflected in reaction times, i.e. do reaction time latencies for the learner groups slowly decrease with growing proficiency? All these questions will be addressed in the discussion of the learner data in comparison with the native speaker data. The results of the two experiments for both learners and natives challenge the common notion of two strictly separated processes when speaking of lexicon and grammar. They also challenge the notion of a strict separation between native speakers and second language learners in terms of processing speed. Rather, all individual groups have their own idiosyncratic behaviour and it can also be assumed that each individual learner behaves in a highly idiosyncratic way, depending on the way he or she has been exposed to the relevant factors influencing processing speed.

The subsequent section will discuss the findings of the lexical decision experiment reported in 7.1. The test design and all methodological considerations are described in detail in section 6.1. In section 8.2 of this final chapter, I will discuss the overall reaction times, the real word effect, the number effect, the interaction between number and absolute dominance, the interaction between number and relative dominance as well as the frequency effect both for the native speaker group and the learner group. In section 8.3, I will discuss the findings of the phrasal grammatical judgment task that were presented in section 7.2. The test design and all methodological considerations in this test are described in section 6.2. I conclude this chapter with some general remarks and suggestions for further research.

8.2 Lexical decision task

8.2.1 overall reaction times

Overall, the native sixth formers responded most quickly to the stimulus words. Their reaction times are faster than those of all learner groups and also faster than those of the native soldiers. In fact, the soldiers were nearly the slowest group, outperformed only by the youngest learner group (7th class), which was slowest of all. The learner groups did not show any regular decrease in reaction time as they got older and had been learning English for more years.

Normally one might predict that reaction times decrease as the learner groups increase in proficiency. This cannot be confirmed by this experiment. The native pupils are indeed fastest but the other groups do not show a regular pattern, except for the youngest group (7th class), which has the least formal language input and slowest reaction times.

One striking finding is the high reaction time of the native soldiers. This may lead to three possible conclusions. A first possible conclusion might be that the methodology was inappropriate. After all, native speakers are believed to have automatised their language. And all native speakers are believed to process language in a qualitatively different way from learners. Learners may never have been exposed to the language in the same way a native speaker typically has. A native speaker therefore has automatised a great deal of his language that the learner has not. Even though automatisation is not the subject of this study and its nature is not fully agreed upon, it is nevertheless automatically implied when comparing the reation times of learners as opposed to those of natives. Why else should one compare natives with learners? Native speakers -independent of educational background, language use, etc.- are regarded as a homogeneous group when they serve as a control group for learner data. So a result that fundamentally challenges this assumption may be based on a methodology that is not sound and should be improved.

The second conclusion one might draw is that reaction times do not tell us what we believe them to tell us. Reaction times may not be a reliable predictor for processing at all. They may not reveal automatised versus unautomatised language processing. And they do not predictably tell us whether someone is a learner or a native speaker.

A third possible conclusion is that learners and native speakers simply cannot be divided into two separate homogeneous groups. As in many other areas, the test situation, the test task, the subjects´ daily routine and other aspects of the subjects´ backgrounds may very well influence their performance on an artifical language test that is not part of their daily routine. Clearly, the soldiers´ native competence is not reflected in their processing speed. This may be because reading words on a computer screen is not part of their daily routine, although it is for secondary school pupils. The amount of exposure to written material may therefore very well be a factor that influences the performance on such a task based on reading. Literacy (in the sense of the amount of daily exposure to written material) and the task itself may be the determining factors responsible for this difference in processing speed between one native speaker group and the other. After all, pupils at a secondary school are used to the fact that they may be asked to respond to a written word in a language class, for instance. Soldiers may also be used to responding but not so much to written prompts. This may mean, then, that not only literacy but also the task itself may play a role in the outcome. If so, difference in reaction times may be an artifact. This explanation reflects some problem in the methodology but at the same time questions the use of lexical decision tasks as a general instrument to test native speakers (or learners) as if they were a homogeneous group. One should therefore always bear in mind that reaction times are not per se an indicator of nativenes or automatisation. And they may not be at all reliable to test language proficiency in learners. These are questions to investigate in future research.

8.2.2 nonce words and real words

All the groups were consistent in showing a difference in reaction time of about 100 ms between the nonce words and the real words; the real words were processed more quickly.

This clear finding suggests that whether the subjects are native speakers or learners, they process real and nonce words at different speeds. Therefore it is a reliable and statistically significant finding for all groups (cf. 6.1.2.2) This unambiguous finding renders this experiment representative since it follows expectations without exception. The results also allow for a probabilistic account of word recognition (Schönefeld, 2001) which predicts that frequency plays a role in processing speed. Nonce words are by nature infrequent and subjects cannot know them in advance. Hence reaction times are longer because they cannot be processed the same way as real words merely on a frequency basis.

Forster´s search model (1976) for lexical access predicts such a difference in processing time between real and nonce words (cf. 3.2.1). The subject must search the access file completely until a word is found. For real words, the search can be aborted as soon as the word is found. For nonce words, the search would continue until all possible access forms are searched and none is found.

This finding does not support the cohort model of lexical access by Marslen-Wilson and Welsh (1978). It predicts that nonce words are identified faster than real words. However, this model is based on spoken word recognition and does not make any claims for visual word recognition (cf. 3.2.4). A phonetically illegal syllable at the beginning of a word can be identified much more quickly as a nonce word than can a phonetically legal syllable for a real word because the latter allows for different possible words. For a visual task, the identification probably takes place after the whole word has been read. So according to the cohort model, the identification times should not differ substantially for written real and nonce words. In this experiment, however, the cohort model does not seem to be a plausible explanation for the reaction times found.

8.2.3 Number effect

The native speakers did not show a significant number effect, even though the soldiers came close to one. The learners showed a significant number effect only for class 7 and class 11. Both groups, processed singulars about 20 ms faster than plurals. No other learner group showed any effect for number.

For the native speakers, both singular and plural forms are sufficiently familiar as to produce no significant difference in processing time between the two. One may assume that both forms are automatised and therefore processing times do not differ very much.

The learner group, however, is not as homogeneous as the natives. The youngest learner group (7th class) does show a clear number effect. In their case one might argue that the relatively infrequent formal language input might lead to such a processing difference between singulars and plurals. They could have had significantly less frequent input of the plural forms of the respective nouns which could therefore cause the processing difference. I have not investigated the formal language input to test this assumption. But given the overall results of the learner groups, it does not seem plausible. If the formal input were the cause of the processing difference, all the younger learner groups should show processing differences up to a certain level. But only class 7 and class 11 show such processing differences between singulars and plurals. This is not a logical sequence in terms of the amount of formal language input obtained and therefore it is hard to make any claims about all the learner groups. It remains interesting to explore why for learners the number factor can play a role in processing whereas it does not for the native speakers.

8.2.4 Interaction number x absolute dominance

The number x dominance interaction is significant for the native pupils with a very strong effect for the plural-dominant nouns. That is, there was a large difference between the relatively low reaction times for plural-dominant nouns in their plural form versus the high reaction times for plural-dominant nouns in their singular form. I also found a strong effect for the singular-dominant words. For the singular forms, it does not seem to make any difference whether the noun is singular or plural dominant. For the plural forms, this difference is large.

The native soldiers do not show a significant interaction effect. For both the singular and the plural-dominant words, they processed the singular forms faster. Again, for the plural forms it does make a bigger difference whether the noun is singular- or plural-dominant than it does for the singular forms.

Most learner groups showed no interaction effect. Only the singular-dominant nouns show different reaction times for singulars and plurals. One learner group (class 9) did display a significant interaction effect. The effect is stronger for the plural forms.

This indicates a qualtitative processing difference between learners and native pupils: first, in the very existence of the interaction effect which is absent for most learners and present for the native pupils, and, second, in the nature of the effect. The native pupils show strong effects for plural dominants, whereas the learner groups show strong effects for singular dominants. The learners may have automatised the singular-dominant words in the singular form just enough that their occurrence in the plural may pose a processing difficulty which reflects in reaction times. This effect is absent in the plural dominants because the learners may not have automatised the plural-dominant plural forms to the same extent as the singular-dominant singular forms. Hence, for the learners, both singular and plural forms in the plural-dominant condition are of comparable processing difficulty. For the native pupils, the effect is clearly visible. It does not matter whether the nouns are singular or plural dominant; both influence the processing speed of either the respective singular forms or the respective plural forms.

The soldiers showed no plural dominant effect. This finding is comparable to that for the German learner groups, who also predominantly show an effect for the singular dominant condition. Again, the different findings for the native soldiers may be the result of their daily routine which does not include reading activities or processing written words to such an extent as for the native pupils. Since the soldiers do, however, show an effect for singular dominants, it is unnecessary to question the whole methodology of the experiment or even the research technique as they do perform as expected in this condition.

Another reason why the methodology of this experiment seems to be sound is that both native groups and some learner groups (at least those that show a significant interaction effect) show a much stronger effect for plural forms as compared to hardly any effect for singular forms. This is exactly as predicted in previous work by Levelt et al (1999), Baayen, Dijkstra & Schreuder (1997), and Baayen et al (2003). Their experiments show that the surface frequency of the words, and not the cumulative frequency of stem- and surface frequency, determines the reaction times for the plural forms, but not for the singular forms. It is exactly this surface frequency that was computed in the CELEX database used for the target words in this experiment. Therefore, the results of this experiment for the groups that show an interaction effect are reliable.

Another very interesting observation is the vast processing differences between different groups of native speakers. It may therefore be justified to question the underlying logic of the initial objective of finding differences between learners and natives. Rather, one might see language processing in either group as a wide continuum on which each speaker finds its own coordinate. This idea is also supported by the fact that one learner group (in the middle of the age range) does in fact show an interaction effect quite like that of the native pupils.

8.2.5 Interaction number x relative dominance

The interaction effect for number and relative dominance applies only to a subgroup of the original stimulus items. I chose only those stimulus words that are highly dominant in either their singular or plural number. This means that the dominance effect I found is actually quite strong and does not allow for all degrees of dominance on the continuum. I expected that the existing effects for absolute dominance would become much stronger with an analysis focussing only on the striking dominance cases.

Indeed, I found a stronger effect for the native pupils, especially for the plural nouns. Here, the interaction effect was again significant for native pupils. Interestingly, the native soldiers also showed a significant interaction effect this time, unlike in the absolute dominance condition. The nature of the effect is similar to that of the previous condition, with a strong effect for singular dominant nouns and a small effect for plural dominant nouns. Important, however, is the statistical significance which was not found in the previous analysis for the absolute dominance values. None of the learner groups showed an interaction effect. Class 12 (the highest class) came close to an interaction effect but it was not significant. Any effect is stronger for the singular-dominant nouns as opposed to the plural-dominant nouns.

The relative-dominance nouns are a subgroup of the absolute-dominance nouns. As expected, the effects seem to increase for both native speaker groups. I take this to mean that the degree of dominance does make a substantial difference in native speaker processing. The matter of dominance is directly linked to probabilistic factors which are relevant for processing. Therefore, as I argued earlier, probabilistic factors are psychologically plausible. My analysis of the two dominance conditions has shown that as soon as the degree of dominance increases, processing changes qualitatively (although it can only be measured in quantitative terms here).

It is nice to see that the change from absolute to relative dominance has no effect for the learners. The learners are not yet prone to such subtle probabilistic factors. They use all their vocabulary comparatively infrequently and cannot automatise it in the same way as the native speakers. In cases where the absolute dominance condition does not yield any interaction effect, a subgroup of the stimulus items is also unlikely to produce any effects. Thus, it is unlikely that it is the smaller size of the subgroup that is responsible for a lack of any effect. Rather, it is the nature of the plural-dominant nouns that may not be plural-dominant in the learner´s language input. One should bear in mind here that singular and plural dominance has been determined by means of a frequency corpus based on native speakers. It is quite natural for language learners to have a completely different probabilistic mindset for their vocabulary.

It is important to note that interaction for relative dominance renders a stronger effect than absolute dominance. That is, the plural or singular dominance is relatively high, compared to absolute dominance in the analysis in 7.1.3.4. This means that the surface frequency of the plural form is substantially different from the stem frequency of the base form. This produces stronger results than the stimulus words that include all levels of dominance. Hay (2001) found this relative frequency to be important for derivational morphemes. This experiment supports Hay´s finding that relative frequencies are more powerful than absolute frequencies and are crucial for processing, even in inflections.

8.2.6 Frequency effect

I found significant frequency effects for all frequency levels in the native pupils. The soldiers show an overall frequency effect. That is, the three frequency levels were different without specifying which levels are significantly different from each other. They did not show differences between high- and mid-frequent nouns, but only between high- and low-frequency and mid- and low-frequency nouns. The frequency effects were significant for all learner groups. Sometimes, the differences between the sublevels were not significant (e.g. I found a significant difference between high- and mid- frequency nouns but not between mid- and low- frequency nouns). For all learner groups, high-frequency nouns stand out and their processing is significantly different from that for all other nouns. In two learner groups (class 7 and 11) there was no difference between mid- and low -frequency nouns but for all other learner groups, this difference was significant. Interestingly, these are the same two learner groups that also show a number effect.

The overall frequency effect is obvious for all groups. This is the predicted outcome based on the literature (Baayen, Dijkstra & Schreuder, 1997; Baayen et al 2003; Münte et al, 1999) which claims that by itself, word stem frequency influences the processing speed of inflected and uninflected words. Therefore, this experiment can be considered a valid one. Frequency seems to have a very strong impact on processing speed.

The frequency effect for native pupils is different at each level and clearly shows the probabilistic balancing between the different levels (based on log-frequencies, considering perceived frequencies). The native soldiers show a difference only between mid- and low-frequency, and high- and low-frequency nouns, but not between high- and mid-frequency nouns. This is not surprising: for native speakers, mid-frequency nouns are still relatively frequent, unlike for learners. Therefore, they may not be salient enough to cause a processing difference for the soldiers, especially since the soldiers are not used to reading and responding to written words in their daily routine. Low-frequency nouns are infrequent enough to result in a processing difference.

The learners show significant overall frequency effects, too. In two cases, the difference between mid- and low-frequency words was not significant (class 7 and 11). Unlike the natives, the learners in these cases may be less familiar with mid-frequency words, and only the high-frequency words occur frequently enough to cause a processing difference. This is especially likely because the word lists were chosen from a native-speaker corpus and the frequencies are native-speaker frequencies. It is impossible to say how often the words were used in the learner input in their foreign language class and in what ways these native-speaker frequencies correspond to language-learner frequencies. The two learner groups that show only the significant effect for the high-frequency words process the mid-frequency and low-frequency words the same; while the other learner groups differentiate significantly between all levels. It would be reasonable to assume that those learners who differentiate between mid- and low levels are more advanced than those that do not show this processing difference. But this is not the case. Class 7 and class 11 do not show this difference in processing. The latter is a relatively high class, second highest in this sample of learners. For class 7 this assumption would hold. Nevertheless, one cannot generalise since the other four learner groups were not consistent in their behaviour. It is interesting to note that the two classes that do not differentiate between mid and low frequencies also show a number effect. As reported in 8.2.3, the number effect in the learner groups could be interpreted as familiarity with singular nouns in the language input of the learners.

The native speaker groups, as well as the other four learner groups (classes 8, 9, 10, and 12) show no number effect. For them, apparantly, plural forms are familiar enough to process like singulars. It is important to note that the difference between singulars and plurals is not a mere frequency difference since I controlled for the frequencies. The list of stimulus words (based on the native speaker corpus) contains equal number of low-frequent plurals and low-frequent singulars. For the learners, however, an occuring number effect might indeed be related to a frequency effect. The language learner might specifically learn the words in the singular form, making the singular form more salient. The data on classes 7 and 11 as opposed to all other subject groups justifies such an assumption. And in line with the lack of any frequency effect between mid- and low-frequency nouns, their overall processing behaviour suggests little language input and a processing stage furthest removed from the native pupils.

In terms of lexical access, the frequency effects found support Morton´s (1979) logogen model of word access. Probabilistic information is crucial in determining processing while accessing a word from the lexicon, exactly what the subjects need to do in this task.

8.3 Phrasal grammatical judgment task

8.3.1 The findings interpreted

The phrasal grammatical judgment task was described in Chapter 8. Subjects had to respond to two-word phrases and decide whether or not the phrase was grammatical. Reaction times were recorded. The aim of this task was to find out whether reaction times differed between two conditions: dual numeral / dual-concept noun and plural numeral / plural noun.

The learners´ reaction times for plural-concept nouns with dual numeral (*two teeth*) were slightly faster than their reaction times for plural-concept nouns with plural numeral (*ten teeth*). This difference was significant only for class 11. The native speakers showed no significant difference between the two conditions, although the native pupils come close to one.

We would not expect a difference in reaction time between the two plural noun conditions. The choice between "two teeth" and "ten teeth" should not matter for processing. Both concepts are equally artificial. However, the results for the plural-concept nouns are not clear. Only one learner group showed a significant reaction time effect, with shorter response latencies for the dual numeral. The other learner groups did not show a significant effect, merely tendencies towards such an effect. This suggests a possible processing advantage for the dual numeral as opposed to the plural numeral in the case of the language learners. This is by no means a real effect if it occurs in only one group and could be considered coincidental. But this tendency among the learners to favour the dual determiner for the plural nouns can also be explained by the "subitising" effect (cf. Dehaene, 2001). Subitising is the term used for the fact that humans can immediately and very accurately name the numerosity of a set of objects smaller than four. This ability is not restricted to a quantity of objects presented; this number size effect can be transferred directly to symbolic number presentation through digits or number words (Dehaene, 2001) and means that people process smaller numbers more quickly than larger numbers. This effect may account for the advantage for the dual numeral as opposed to the plural numeral (> 2) in the phrasal grammatical judgement task. Therefore it may be an artefact, not at all related to the conceptual relationship between numeral and noun. Since the native speakers and most learner groups did not differ significantly between the two conditions, we should not assume a general subitizing effect in this experiment. This is merely an attempt to explain the behaviour of one learner group.

For the dual concept nouns, the results are more interesting. Both native speaker groups performed significantly more quickly in the dual numeral / dual-concept noun condition (*two eyes*) than in the plural numeral / dual-concept noun condition (*ten eyes*). Among the learners, those in classes 7 and 11 showed the same result. Class 9 came close to showing a significant difference between the two conditions and classes 10 and 12 did not show a significant difference between the two sets of stimuli. In 6.2.3. I explained why I lack the data of class 8 for this experiment.

For both native speaker groups, this is a clear result. It is indeed faster to process dual-concept determiners with dual-concept nouns than plural numerals and plural-concept nouns. This means that the conceptual level does influence the processing of pluralised noun phrases. Grammatical agreement alone does not determine the processing speed; conceptual information is also a strong determinant for native language processing.

In case of the learners, this conclusion does not hold without exceptions. Two of the learner groups (classes 7, and 11) did behave like the native speakers, but three learner groups (classes 9, 10 and 12) did not significantly differ between the two conditions, although one came close to a significant effect (class 9). Again, the level of the learner groups in terms of age and years of English taught at school does not predict the outcome. The two groups showing native-like behaviour are not the two top level groups and the two groups not showing any effect are not the bottom-level groups. So language development towards native-speaker processing behaviour is not at all related to one´s class in school. Why do native speakers so clearly show an effect here and learners a less clearly one? It could be explained by the degree of language proficiency that determines whether or not conceptual information may enter the processing level. The very fact that we are talking about conceptual information here should suggest that it is language independent and the effect should be visible in all subject groups. However, the learners might possess a language proficiency threshold which allows them to acknowledge or ban all probabilistic information (including the fact that dual determiners very often co-occur with dual concept nouns) in processing. This information should be present in their native language but simply using a foreign language already implies a high processing load. Such subtle information as probabilistic factors could benefit relatively little in processing time compared to the processing time load of words learned by rote. So, compared to the processing load learners already carry, the probabilistic factors might play only a minor role. The fact that some learners do show an effect means that they are indeed sensitive to probabilistic and conceptual factors but the normal tasks involved in processing a foreign language outweigh the language-independent factors.

8.4 General remarks

8.4.1 Suggestions for further research

This work has provided some interesting findings that shed new light on the nature of language processing among learners and native speakers. Many of these findings, however, provide an impetus for further research since many questions remain unanswered. In addition, the literature reviewed in the first five chapters provides ample ideas for new research when applied to the conditions described in this thesis.

For instance the experiment by Münte et al (1999), described in section 5.3, looked at processing differences between regular and irregular verbs in a priming experiment. Interestingly, they found that priming effects cannot be measured using reaction times; instead, they found that different types of brain activity occurred in the ERP condition. Having ruled out formal aspects that could have caused such effects in the control condition, they concluded that this must be a morphological effect. Such a basic morphological effect should also be found in noun inflections. It would be extremely interesting to see if these effects also occur in nouns, and whether the different subject groups respond differently to the nouns in terms of brain activity. Having looked at reaction times in my experiments, it would now be interesting to see, what the processing looks like from a qualitative point of view, namely in brain activity. In a subsequent study, all my experiments could be repeated while measuring event-related brain potentials.

The findings of my experiments reported in Chapter 7 yield enough data for several follow-up experiments. One striking observation is the apparent similarity in the way the learner groups (classes 7 and 11) behave. However wide apart they may be in terms of formal language input, they seem to be similar in the number effect they show, unlike the other learners or the native speakers. In this respect, they seem to be the least advanced in processing terms, since they process singulars more quickly than plurals. None of the other groups showed such a difference and none had difficulty processing plurals. In addition to the number effect, the frequency effect is also similar in those two groups. All the subject groups show a frequency effect but only in classes 7 and 11 was there no significant difference between the mid- and the low-frequency condition (cf 8.2.6). Apparently, the pupils in classes 7 and 11, are not familiar enough with the mid-frequency nouns to process them differently from low-frequency nouns; in contrast, all the other learners and the natives can differentiate between the two in processing terms. Again, this would suggest that those two groups show the least advanced processing-behaviour.

It is also striking that classes 7 and 11 behave so similarly. First of all, this finding is interesting because it seems to contradict those for all the other learner groups. Second, it would suggest that the way the two groups process these words is task independent if they show such a consistent pattern throughout the experiments. But the most striking finding of all is the outcome in the phrasal grammatical judgment task. Here, classes 7 and 11 show the same effects as both the native speaker groups and are unlike all the other learner groups. In this experiment, the behaviour of these two groups seems to be native-like and thus more advanced than the other learner groups - quite the reverse of what was observed in the first experiment. This certainly is a finding that deserves to be looked at in closer detail because it is exactly this pair of learner groups that agrees on a processing pattern that is different from that of the other learner groups.

In addition, one could look at many more learner groups of different levels to search for a connection between the different test tasks and the processing behaviour of the groups. These findings raise many questions. For example, why is it that formal language input does not seem to predict the processing outcome of a learner group? And why is it that in some tasks, classes 7 and 11 seem to behave far from native-like whereas in the phrasal grammatical judgment task, they do exactly what both groups of native speakers do? Given this finding, it would be interesting to find the common denominator for classes 7 and 11: What is it that makes them process nouns so similarly if not the formal language input? These issues need to be looked at more closely by repeating this series of tasks with many different subject groups.

Another issue to be explored in detail in a follow-up study is the degree of dominance that seems to play a role in the interaction effect. Since the outcomes differed with respect to absolute or relative dominance, it is likely that different findings can be provoked by experimenting with the degree of dominance. This degree of dominance can vary arbitrarily. It would be interesting to determine at what point dominance seems to play a role for processing, i.e. to find out what dominance really is in processing terms. As has been seen in this experiment, this is different for each group. Perhaps there is a tipping point at which one can clearly say that a dominance level is relevant only for native speakers.

Another important finding of this work is the seeming relativity when it comes to native speakers opposed to non-native speakers. If anything, this work has shown that most conditions in the experiments did not yield consistent results for the two native-speaker groups. This was certainly the case for the learners but also -and unexpectedly so- for the natives. Further research should be conducted to re-define the role of the native speaker when it comes to processing. Perhaps a brain imaging study can offer a solution; it might be possible to detect a different quality in brain activity rather than a different quantity (reaction times) in processing between native and non-native groups. From the reaction-time experiments here, it is impossible to conclude that there is a strict division between native and non-native speakers.

8.4.2 Concluding remarks

From all the findings discussed above, it seems clear that extra-linguistic factors play a role in processing. This has been shown for probabilistic factors such as frequency and dominance and for conceptual factors such as number and agreement between determiner and noun. Therefore, it seems, processing does not only distinguish between retrieval from rote memory and retrieval from grammatical processor when the person is applying a rule. This means that the plural inflection in nouns is not merely a matter of grammatical processing and that the singular nouns are not merely stored in rote memory. The plural suffix is not processed exclusively by computation; rather, the degree of “plurality” determines the processing route. The strict separation between grammar and lexicon cannot be sustained. The linguistic number system (singular versus plural) does not denote conceptual numerosity by itself; the plural use in nouns also implies the use of probabilistic information that goes beyond simply adding a plural suffix by applying a rule. The regular -s plural is not such a simple plural form after all; there is much more to it than the plural rule we find in grammar books.

A second important observation can be made in comparing the native speakers with the language learners. There seems to be no regular develomental hierarchy for the learner groups in terms of plural processing. One explanation could be that their exposure to the English language does not regularly increase from one class to the next. If that were happening, it would presumably show up in processing terms. The learners seem to get sufficient input from outside the classroom to blur any clear developmental dimension in processing from one class to the next. For teachers, this is a sobering observation: in processing terms, there is no real development. This also makes it impossible to pinpoint exactly when a learner achieves native-like processing skills. At least this does not appear to be possible during a learner´s German *Gymnasium* career.

At this point it should be possible to make a statement about native language processing. But even here, there are vast differences. The native speaker group is not uniform. The above experiments underline the importance of daily routine for good results on artificial tasks like those in the experiment. If we recall the overall reaction times, the two native speaker groups were almost at extreme ends of the continuum (cf 7.1.4). But it is exactly in the processing that one would expect to find a common native speaker who is clearly discernible from a language learner. This is not so much the case for real language use, since it can vary so much from region to region or speaker to speaker, but one could argue that processing should define the native speaker through the automatisation of language structures. But my experiments challenged exactly this idea. In processing terms, there may be no such thing as a “common” native speaker and a “common” language learner. Based on these results, I would argue instead for a processing continuum for all individuals irrespective of age or school year.

This may come as a relief after the sobering observation for teachers that no developmental dimension exists. The older and more experienced school children were not appreciably faster on the tests than the younger ones and they did not consistently improve their processing from one year to another, but in fact processing times may not actually be a clear indicator of language development. Apparently, so many other factors influence processing, including even the subjects´ daily routines, that it is hard to make a one-dimensional claim in terms of processing development based only on reaction times. Instead, language processing is a multi-dimensional event which includes linguistic and extra-linguistic factors.

The implications for teaching should therefore be clear. Language classes should provide a multi-dimensional approach to learning. Teaching should not only focus on rote learning and rule application; it should also make use of such facilitative factors as frequency and dominance of words. New vocabulary in school books could therefore be based on native speaker corpora that include frequency information. The power of the frequency effect does, in fact, support the idea of repetition. I would not, however, plead for behavioristic repetition of structures; rather, I believe we need meaningful repetition and consolidation of new vocabulary or structures.

None of this is entirely new, of course. A key finding is that native speakers are not a uniform group and it is therefore not necessary to achieve a native-speaker competence in processing terms since there seems to be no such thing. Nor can we point to any dimensional aspect in processing. Learners might learn and also lose vocabulary over and over again. They might be susceptible to factors such as frequency or dominance at different points in their learning career. Even though the frequency variable is a factor based on native speaker corpora in this experiment, learners did seem to be sensitive to it. They are aware of native speaker frequencies. Nevertheless, their reaction times are not comparable to those of native speakers and do not show a developmental dimension. Learning is very idiosyncratic.

Summing up all the above conlcusions, I would plead for a processing continuum which allows various degrees of processing for learner groups and individual learners. This continuum is a multi-dimensional co-ordinate which relies on linguistic, probabilistic, and conceptional factors, and also crucially, on the daily routine of the language user. Processing cannot be reduced to a simplistic event that readily tells us something about a proficiency level.

Bibliography

Anderson, S.R. (1992). *A-Morphous Morphology.* Cambridge Studies in Linguistics 62: Cambridge University Press.

Baayen, R. H., Dijkstra, T. & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. In: *Journal of Memory and Language* 37, 94-117.

Baayen, R. H. (2003). Probabilistic Approaches to Morphology. In: *Probabilistic Linguistics.* Cambridge (Mass.): MIT Press.

Baayen, R. H., McQueen, J., Dijkstra, T. & Schreuder, R. (2003). Frequency effects in regular inflectional morphology: Revisiting Dutch plurals. In Baayen, R.H. and Schreuder, R. (eds), *Morphological structure in language processing*, Berlin: Mouton de Gruyter, 355-390.

Baayen, R. H., Piepenbrock, R., & Gulikers L. (1995). *The CELEX Lexical Database* (CD-Rom). University of Pennsylvania, Philadelphia, PA: Linguistic Data Consortium.

Badecker, W. & Caramazza, A. (1989). A lexical distinction between inflection and derivation. *Linguistic Inquiry, 20*, 108-116.

Bates, E. & Goodman, J. (1997). On the inseparability of grammar and lexicon: evidence from acquisition, aphasia and real-time processing. In: *Language and Cognitive Processes* 12 (5/6), 507-584.

Beard, R. (1982). The plural as a lexical derivation. In *Glossa* 16:2.

Berko, J. (1958). The child´s learning of English morphology. In: *Word* 14, 15-177.

Berretta, M. (1995). Morphological markedness in L2 acquisition. R. Simone (ed). *Iconicity in Language*. Amsterdam, Benjamins, 197-233.

Bertram, R. et al. (2000). Affixal Homonymy triggers full-form storage, even with inflected words, even in a morphologically rich language. In: *Cognition* 74, B13-B25.

Bod, R., Hay, J. & Jannedy, S. (2003). Introduction. In: *Probabilistic Linguistics.* Cambridge (Mass.): MIT Press.

Booij, G. (1993). Against split morphology. In: Booij, G. & van Marle, J. (eds), *Yearbook of Morphology,* 27 - 49. Dordrecht: Kluwer Academic Press.

Broersma, M. & de Bot, K. (2001) De triggertheorie voor codewisseling. In: *Toegepaste Taalwetenschap in Artikelen 65*. ANeLA: 41-45.

Bowers, J. S., Vigliocco, G., Stadthagen-Gonzalez, H. & Vinson, D. (1999). Distinguishing language from thought: experimental evidence that syntax is lexically rather than conceptually represented. In: *Psychological Science*, 10 (4): 310-315.

CELEX lexical database homepage (2001). www.let.kun.nl/celex/

Cenoz, J. (2002). The development of cross-linguistic influence in third language acquisition: Implications for the organization of the multilingual mental lexicon. Basel: EuroSLA12, 19. September 2002.

Chen, H.-C. & Ho C. (1986). Development of stroop interference in Chinese- English bilinguals. In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15. 316-325.

Comrie, B. (1989). *Language Universals and Linguistic Typology*. Oxford: Blackwell.

Contini-Morava, E. (2000). Noun class as number in Swahili. In: *Between Grammar and Lexicon.* Amsterdam: John Benjamins.

Croft, W. (1990). *Typology and Universals*. Cambridge: Cambridge Univeristy Press.

Davis, M.H., van Casteren, M. & Marslen-Wilson, W.D. (2003). Frequency effects in processing inflected Dutch nouns: a distributed connectionist account. In: H. Baayen & R. Schreuder (eds.). *Morphological Structure in Language Processing.* Berlin-New York: Mouton de Gruyter.

de Bot, K. (1992). A bilingual production model: Levelt's Speaking Model Adapted. In: *Applied Linguistics* 13.

de Groot, A. M. B. (1993). Word-Type Effects in Bilingual Processing Tasks. In: Schreuder & Weltens (eds). *The Bilingual Lexcion*. Amsterdam: John Benjamins.

Deacon, T. W. (2000). Beneath Linguistics: *Language as a Complex Coevolutionary Phenomenon*. MPI Nijmegen Lectures, December 18.-20. 2000.

Dehaene, S. (2001). Précis of the number sense. In: *Mind & Language*, 16, 16-36.

Emmorey, K. (1995). Processing the dynamic visual-spatial morphology of signed languages. In: L. B. Feldman (ed.). *Morphological Aspects of Language Processing*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Felser, C. (2001). The processing of ambiguous sentences by first and second language learners of English. Paderborn: Interdisziplinaeres Kolloquium, 28. November 2001.

Forster, K.I. (1976). Accessing the mental lexicon. In Wales and Walker (eds.) 1976.

Frijn, J. & de Haan, G. (1994). *Het Taallerend Kind.* Dordrecht: ICG Publications.

Fromkin, V. (1971). The non-autonomous nature of anomalous utterances. In: *Language* 47: 27-52.

Garman, M. (1990). *Psycholinguistics.* Cambridge: CUP.

Giacalone Ramat, A. & Klein, W. (2000). Some questions concerning the development of inflectional morphology. Report of the Euroconference on the Structure of Learner Languages.

Goorhuis, S. M. & Schaerlaekens A. M. (1994). *Handboek Taalontwikkeling, Taalpathologie en Taaltherapie bij Nederlandssprekende Kinderen*. Utrecht: De Tijdstroom.

Green, D. W. (1993). Towards a Model of L2 Comprehension and Production. In: R. Schreuder & Weltens (eds). *The Bilingual Lexcion*. Amsterdam: John Benjamins.

Grosjean, F. (1980). Spoken word recognition and the gating paradigm. In *Perception & Psychophysics*, 28, 267-283.

Hakansson, G. (2003). Language, Gestures and Pictures in the Perspective of Semiotic Development. Paper at the European workshop on language learning at the University of Paderborn in April 2003.

Halgren, E., Wang, Ch., Ulbert, I., Knake, S., Marinkovic, K., Wu, J. & Schomer, D.L. (2005). Processing stages underlying word recognition in the anteroventral temporal lobe. (in press).

Hancin-Bhatt, B. & Nagy, W. (1994). Lexical transfer and second language morphological development. In *Applied Psycholinguistics* 15: 289-310.

Hay, J. (2001). Lexical frequency in morphology: is everything relative?. In: *Linguistics* 39 (6), 1041-1070. Berlin and New York: de Gruyter.

Hirtle, W. (2000). Number in the English Substantive. In: E. Contini-Morava & Y. Tobin (eds). *Between Grammar and Lexicon*. Amsterdam: John Benjamins.

Holmes, F. (1996). Cross-Language Interference in Lexical Decision. Internet publication: www.phon.ucl.ac.uk

Hurford, J. (1987). *Language and Number*. Oxford: Blackwell.

Janda, L. A. (2000). From Number to Gender, from Dual to Virile. In: E. Contini-Morava & Y. Tobin (eds). *Between Grammar and Lexicon.* Amsterdam: John Benjamins.

Janssen, D., Roelofs, A. & Levelt, W.J.M. (2002). Inflectional frames in language production. In: *Language and Cognitive Processes* 17 (3): 2009-236.

Janssen, D. (1999). *Producing Past and Plural Inflections.* Nijmegen: MPI Series in Psycholinguistics, 11.

Jin, Y.S. (1990). Effects of concreteness on cross-language priming in lexical decisions. In: *Perceptual and Motor Skills* 70. 1139-1154.

Jones, N. (2004). The Neurobiology of memory consolidation. In: *The Neurobiology of Language: Perspectives from Second Language Acquisition.* 111-142. J. Schumann, S. Crowell, N. Jones, et al (eds.). New Jersey: Lawrence Erlbaum Associates.

Kirsner, K. et al. (1984). The Bilingual Lexicon: Language-specific units in an integrated network. In: *Journal of Verbal Learning and Verbal Behavior* 23, 519-539.

Klein, W. (2000). Inflectional morphology and second language acquisition. Discussion paper for the workshop „The dynamics of learner varieties“ in San Feliu.

Klein-Gunnewiek, L. (2000). *Sequenzen und Konsequenzen*. Amsterdam: Rodopi B.V.

Krashen, S. (1983). *Second Language Acquisition and Second Language Learning*. Oxford: Pergamon.

Kroll, J. F. (1993). Accessing Conceptual Representations for Words in a Second Language. In: R. Schreuder & B. Weltens (eds). *The Bilingual Lexcion*. Amsterdam: John Benjamins.

Kroll, J. F. & Curley, J. (1988). Lexical memory in novice bilinguals: The role of concepts in retrieving second language words. In: *Practical Aspects of Memory: Current Research and Issues*, Vol. 2 ed. by M.M. Gruneberg, P.E. Morris & R.N. Sykes, 389-395. London: John Wiley & Sons.

Kupfermann, I. (1985). Learning. In: E. R. Kandel & James H. Schwartz (eds.). *Principles of Neural Science*, 805-815. New York: Elsevier.

Laufer, B. (1989). A factor of difficulty in vocabulary learning: deceptive transparency. In: Nation & Carter (eds), 1989. AILA Review 6 / 1989, 10-20.

Levelt, W.J.M. (1989). *Speaking: From intention to articulation.* MIT Press.

Levelt, W.J.M. (1993). Language use in normal speech and its disorders. In: G. Blanken, J. Dittmann, H. Grimm, J. Marshall & C. Wallesch (eds). *Linguistic Disorders and Pathologies: An International Handbook*. Berlin: De Gruyter, 1-15.

Levelt, W.J.M. (1995). The ability to speak: from intentions to spoken words. In: European Review 3 (1): 13-23.

Levelt, W.J.M., Roelofs, A. & Meyer, A.S. (1999). A theory of lexical access in speech production. In *Behavioral and Brain Sciences* 22 (1): 1-75.

Logan, G.D. (1988). Toward an instance theory of automatization. *Psychological Review*, 95: 492-527.

Lück, M., Hahne, A. & Clahsen, H. (2001). Dissociations of regular and irregular German plurals in an auditory ERP study. 23. Jahrestagung der deutschen Gesellschaft fuer Sprachwissenschaft: Leipzig, 28. February, 2001.

Mägiste, E. (1984). Stroop tasks and dichotic translation: The development of interference patterns in bilinguals. In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 10. 304-315.

Manning, Ch. D. (2003). Probabilistic Syntax. In: *Probabilistic Linguistics*. Cambridge (Mass.): MIT Press.

Marinis, T. (2003). Psycholinguistic techniques in second language acquisition research. In: *Second Language Research* 19,2: 144-161.

Marinkovic, K. (2004). Spatiotemporal dynamics of word processing in the human cortex. In: *The Neuroscientist*: 142-152.

Marslen-Wilson, W. & Welsh, A. (1978). Processing interactions and lexical access during word-recognition in continuous speech. *Cognitive Psychology* 10: 29-63.

McWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk.* Mahwah, NJ: Lawrence Erlbaum Associates (3rd ed.).

Meara, P. (1984). The study of lexis in interlanguage. In: *Interlanguage*, A. Davies, C. Criper & A.P.R. Howatt (eds.). Edinburgh: Edinburgh University Press.

Meyer, D. E. & Schvaneveldt, R. W. (1971). Facilitation in recognising pairs of words: Evidence of a dependence upon retrieval operations. In *Journal of Experimental Psychology, 90*: 227-234.

Morton, J. (1969). Interaction of information in word recognition. In *Psychological Review* 76: 165-178.

Morton, J. (1979). Word recognition. In: *Psycholinguistics series 2: structures and processes*, P. Kolers, M. Wrolstad & H. Bouma (eds.). New York: Plenum.

Moyer, R.S. & Landauer, T. K. (1967). Time required for judgments of numerical inequality. *Nature 215:* pp. 1519-1520.

Münte, T. F., Say, T., Clahsen, H., Schlitz, K., & Kutas, M. (1999). Decomposition of morphologically comlex words in English: evidence from event-related brain potentials. In *Cognitive Brain Research* 7: 241-253. New York: Elsevier.

Neef, M. (2001). A non-universal approach to English inflectional morphology. In *Linguistics 39 (6)*: 1149-1170. Berlin and New York: de Gruyter.

Pavlenko, A. (1999). New approaches to concepts in bilingual memory. In: *Bilingualism: Language and Cognition* 2 (3), 209 - 230.

Perea, M., & Lupker, S.J. (2003). Transposed letter confusability effects in masked form priming. In: S. Kinoshita & S.J. Lupker (Eds.), *Masked priming: State of the Art* (pp.97-120). Hove, UK: Psychology Press.

Perlmutter, D.M. (1988). The split morphology hypothesis. In M. Hammond & M. Noonan (Eds.), Theoretical morphology: approaches in modern linguistics (pp. 79-100). San Diego, CA: Academic Press.

Pienemann, M. (1998). *Language Processing and Second Language Development: Processability Theory.* Amsterdam: John Benjamins.

Pinker, S. (1999). *Words and Rules: The Ingredients of Language*. New York: Harper Collins.

Poelmans, P. (2003). *Developing Second-Language Listening Comprehension*. PhD Thesis, Utrecht: LOT.

Rawlinson, G.E. (1976). *The Singnificance of Letter Position in Word Recognition*. Unpublished PhD Thesis, Psychology Department, University of Nottingham, Nottingham UK.

Rohdenburg, G. (2004). Variable plural marking with measure nouns in non-standard English and low German dialects. (in press).

Sabourin, L. (2001). L1 effects on the processing of grammatical gender in L2. In: S. Foster-Cohen & A. Nizegorodcew (eds.) *EUROSLA yearbook. Volume I, 2001.* Amsterdam: John Benjamins.

Sasse, H.J. (1993). Syntactic categories and subcategories. In: J. Jacobs, A. von Stechow, W. Sternefeld & T. Vennemann (eds.), *Syntax. Ein Internationales Handbuch zeitgenössischer Forschung / An International Handbook of Contemporary Research*, 646-686. Berlin: Mouton de Gruyter.

Say, T. & Clahsen, H. (1999). Words, rules and stems in the Italian mental lexicon. *Essex Research Reports in Linguistics* 27, 2-47.

Schönefeld, D. (2001). *Where Lexicon and Syntax Meet.* Berlin: Mouton de Gruyter.

Schriefers, H. (1993). Syntactic Processes in the Production of Noun Phrases. In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19 (4), 841-850.

Schriefers, H., Friederici, A. & Graetz, P. (1992) Inflectional and derivational morphology in the mental lexicon: Symmetries and assymmetries in repetition priming. *The Quarterly Journal of Experimental Psychology, 44A*, 373-390.

Segalowitz, N. & Segalowitz, S. (1993). Skilled performance, practice, and the differentiation of speed-up from automatization effects: Evidence from second language word recognition. *Applied Linguistics* 14, 369-385.

Sereno, J. A. & Jongman, A. (1997). Processing of English inflectional morphology. In *Memory & Cognition* 25 (4), 425-437.

Shattuck-Hufnagel, S. (1983). Sublexical units and suprasegmental structure in speech production planning. In: P. F. Mac Neilage (ed.) *The Prodcution of Speech*. New York: Springer-Verlag.

Singleton, D. (1999). *Exploring the Second Language Mental Lexcion*. Cambridge: Cambridge University Press.

Smith, P.T., Meredith, T, Pattison, H.M. & Sterling, C. (1984). The representation of internal word structure in English. In: Henderson (ed.) 1984.

Sonnenstuhl, I., Eisenbeiss, S. & Clahsen, H. (1999). Morphological priming in the German mental lexicon. In: *Cognition* 72, 203-236.

Stowe, L. & Kaan, E. (2001). *Developing an Experiment: Techniques and Design*. Klapper voor het werkcollege Taalkunde LTA50DO6. Rijksuniversiteit Groningen.

Svensson, P. (1998). *Number and Countability in English Nouns: an Embodied Model.* Uppsala: Swedish Science Press.

Taft, M. (1981). Prefix strippings revisited. In: *Journal of Verbal Learning and Verbal Behaviour* 20: 289-297.

Taft, M. & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. In: *Journal of Verbal Learning and Verbal Behaviour* 14: 638-647.

Talamas, A., Kroll, J. F. & Dufour, R. (1999). From form to meaning: Stages in the acquisition of second language vocabulary. In: *Bilingualism: Language and Cognition* 2 (1), 45-58. Cambridge University Press.

Tobin, Y. (2000). The dual number in Hebrew. In: E. Contini-Morava & Y. Tobin (eds). *Between Grammar and Lexicon.* Amsterdam: John Benjamins.

Tobin, Y. (1990). *Semiotics and Linguistics.* New York: Longman.

Tomasello, M. (2002). *Evolutionary Psychology and Primate Cognition*. MPI Nijmegen Lectures, December 11-13.

Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Cambridge MA: Harvard University Press.

Tomasello, M. (2000). Do young children have adult syntactic competence?. In: *Cognition 74*, 209-253.

Tuggy, D. (1997). On the storage vs. computation of complex linguistic structures. Unpublished conference paper from the ICLC Amsterdam.

Ullman, M. T. (2004). Contributions of memory circuits to language: the declarative/ procedural model. In: *Cognition 92*, 231-270.

Ullman, M. T., Newman, A., Izvorski, R., & Neville, H. (2000). Electrophysiological correlates of lexicon and grammar. Online abstract of a paper presented at the CUNY Sentence Processing Conference, San Diego.

van de Craats, I. (2002). Taal(on)bewustzijn en tweede-taalverwerving. *In: Toegepaste Taalwetenschap in Artikelen* 68, 27-37.

van der Lugt, A. (1999). *From Speech to Words.* Nijmegen: MPI Series in Psycholinguistics, 13.

Vigliocco, G., Lauer, M., Damian, M.F. & Levelt, W.J.M. (2002). Semantic and syntactic forces in noun phrase production. In *Journal of Experimental Psychology: Human Perception and Performance 28*, 46-58.

Vigliocco, G. (2000). Language processing: the anatomy of meaning and syntax. In Current Biology 10, R78-R80.

Vigliocco, G., Vinson, D., Martin, R. & Garrett, M.F. (1999). Is “count” and “mass” information available when the noun is not? An investigation of tip of the tongue states and anomia. In *Journal of Memory and Language* 40, 534-558.

von Humboldt, W. (1836). Ueber die Verschiedenheit des menschlichen Sprachbaus und ihren Einfluss auf die geistige Entwicklung des Menschengeschlechts. In: P. H. Salus (ed.) *On Language: Plato to von Humboldt*. New York: Holt, Rinehart and Winston, Inc.

Wartheimer, M. (1938). Laes of organisation in perceptual forms. In W. D. Ellis (ed.), *A Source Book of Gestalt Psychology*. London: Routledge & Kegan, 71-94.

Weinreich, U. (1953). *Languages in Contact: Findings and Problems.* New York: Linguistic Circle of New York. Reprinted in 1974 by Mouton, The Hague.

Whalen, J., Gallistel, C. R. & Gelman, R. (1999). Nonverbal counting in Humans: The psychophysics of number representation. In: *Psychological Science*: 10, 2 (130-137).

Appendix A

The instructions presented on the computer screen before the start of the lexical decision experiment:

Welcome to the experiment!

You will see a string of letters on the screen.

Your task is to determine whether the target word

is an existing word or a non-word.

Press the red key for a non-word and

the green key for a real word.

You may request the following

item by pressing the spacebar.

PRESS THE SPACEBAR TO BEGIN!

The instructions presented on the computer screen before the start of the phrasal grammatical judgment experiment:

> Hello and welcome to the experiment!
>
> You will see two words on the screen. Your task is to determine whether this is a correct phrase or not. So, for instance, if you see “she talk”, it is not correct because it lacks an “s”in “talk”. A phrase like “two door” is also incorrect because it lacks an “s”. But “the babies” would be correct.
>
> Press the green key for a correct phrase and the red key for an incorrect phrase.
>
> You will go through 10 practice items before the real test starts.
>
> Press the spacebar to begin!

For both experiments, the response keys on the keyboard are marked with red and green tape in order to facilitate quick responses. The subjects are asked to use both hands and keep their two index fingers on top of the keys during the whole experiment in order to respond as fast as possible. The “V” key on the German keyboard is marked as the red key and the “N” key is marked as the green key. These response keys were chosen because they allow a comfortable position for both hands while taking the 30-minute test. The key on the right hand side was chosen to be the response key for the correct words because most participants are right handed and will therfore have a right-hand advantage. Since all target stimulus words are real words, they should all be responded to with the green right hand key for “real word”.

Appendix B

A list of all the 385 stimulus items of the lexical decision task in the order of one random test run (including ten trial items):

tog (nonce), wan (nonce), tegs (nonce), droab (nonce), cloam (nonce), shown (real), took (real), pamced (nonce), twens (real), learp (nonce), drawn (real), bloody (real), pity (real), gave (real), plenty (real), loaked (nonce), does (real), bother (real), all (real), blown (real), ticrac (nonce), thligs (nonce), veag (nonce), yee (nonce), mild (real), ournd (nonce), wull (nonce), eggs (real), off (real), lalt (nonce), carled (nonce), nat (nonce), frund (nonce), years (real), sache (nonce), negtle (nonce), join (real), riel (nonce), fool (real), shot (real), hair (real), gets (real), as (real), az (nonce), shown (real), touh (nonce), bowed (real), lian (nonce), tear (real), wrand (nonce), dudley (nonce), eye (real), waits (real), frew (nonce), keys (real), deadly (real), damn (real), kust (nonce), herbs (real), dack (nonce), dahn (nonce), rheb (nonce), loof (nonce), rowd (nonce), boak (nonce), wobed (nonce), reasp (nonce), oet (nonce), soume (nonce), bluck (nonce), dat (nonce), hot (real), taled (nonce), learp (nonce), hily (nonce), rosh (nonce), ipe (nonce), fell (real), pell (nonce), goined (nonce), naut (nonce), treamd (nonce), grof (nonce), ruins (real), spies (real), geg (nonce), snake (real), eeb (nonce), kibe (nonce), doobly (nonce), booking (real), keen (real), flew (real), pear (real), coul (nonce), suirn (nonce), birds (real), looks (real), hud (nonce), stuck (real), talked (real), sieda (nonce), rofty (nonce), big (real), droab (nonce), rud (nonce), rewth (nonce), locder (nonce), rails (real), took (real), lod (nonce), foot (real), mand (nonce), tog (nonce), month (real), trigh (nonce), chald (nonce), jien (nonce), coet (nonce), knoobig (nonce), gained (real), aimed (real), rocky (real), pamced (nonce), slup (nonce), discs (real), not (real), ill (real), twens (real), looked (real), wenst (nonce), arts (real), swait (nonce), doed (nonce), udd (nonce), poad (nonce), hard (real), olives (real), miles (real), black (real), used (real), near (real), made (real), banged (real), can´t (real), words (real), dirb (nonce), toag (nonce), silly (real), fed (real), lisly (nonce), sealt (nonce), meac (nonce), wain (nonce), dod (nonce), weepy (real), up (real), chips (real), lowb (nonce), toob (nonce), own (real), ruced (nonce), seud (nonce), luibt (nonce), old (real), ducks (real), tipy (nonce), dreamt (real), pounds (real), picked (real), schools (real), asks (real), nees (nonce), low (real), stoet (nonce), kid (real), gib (nonce), knees (real), scroaned (nonce), hichp (nonce), limd (nonce), tunt (nonce), oivel (nonce), knucked (nonce), red (real), day (real), nat (nonce), camped (real), oact (nonce), syp (nonce), faced (real), deag (nonce), add (real), ricep (nonce), far (real), daiged (nonce), wried (nonce), dusc (nonce), rhad (nonce), takes (real), rumd (nonce), friends (real), skas (nonce), kates (nonce), bees (real), built (real), gid (nonce), wilked (nonce), men (real), able (real), shead (nonce), beal (nonce), plus (real), cat (real), grew (real), hands (real), goat (real), rilg (nonce), tewn (nonce), here (real), reat (nonce), plugged (real), rean (nonce), urt (nonce), legs (real), lil (nonce), mouse (real), fid (nonce), asode (nonce), truths (real), teab (nonce), rouh (nonce), raih (nonce), calmed (real), truned (nonce), walked (real), rights (real), soeg (nonce), aref (nonce), sold (real), brought (real), had (real), colder (real), pills (real), hos (nonce), aren´t (real), lep (nonce), goes (real), tay (nonce), arctic (real), weird (real), malp (nonce), croy (nonce), knalb (nonce), thober (nonce), boot (real), laifed (nonce), dirty (real), ants (real), early (real), dives (real), books (real), bikes (real), pewy (nonce), ayw (nonce), aged (real), nus (nonce), wonsh (nonce), cloudy (real), tents (real), drum (real), possed (nonce), tegs (nonce), came (real), laery (nonce), knocked (real), nact (nonce), nakes (nonce), foced (nonce), gentle (real), called (real), marr (nonce), boat (real), nail (real), nomth (nonce), pea (real), thiff (nonce), got (real), kipsed (nonce), turned (real), seen (real), bill (real), thought (real), sild (nonce), aside (real), round (real), wreg (nonce), glupped (nonce), thoun (nonce), bat (real), toof (nonce), cint (nonce), ald (nonce), laid (real), heads (real), died (real), soed (nonce), screened (real), way (real), child (real), claim (real), forty (real), aches (real), works (real), slight (real), wroks (nonce), balls (real), foff (nonce), laed (nonce), hour (real), tough (real), lig (nonce), wol (nonce), raf (nonce), blank (real), fifth (real), telnpy (nonce), earth (real), pissed (real), turth (nonce), necks (real), fare (real), kneg (nonce), lelf (nonce), arm (real), dental (real), min (nonce), sield (nonce), doulcy (nonce), slide (real), slooch (nonce), rogthub (nonce), girl (real), yare (nonce), threw (real), cilmed (nonce), tree (real), eyk (nonce), holy (real), telked (nonce), tendal (nonce), failed (real), aunts (real), palm (real), pearls (real), frog (real), trips (real), niek (nonce), well (real), aep (nonce), prices (real), toe (real), hois (nonce), moli (nonce), went (real), did (real), dealt (real), sun (real), out (real), rhee (nonce), tenar (nonce), tridy (nonce), reet (nonce), broad (real), cloam (nonce), meid (nonce), toast (real), vides (nonce), imead (nonce), neek (nonce), prit (nonce), heam (nonce), aimed (real), shoe (real), least (real)

Deutsche Zusammenfassung

Ziel der vorliegenden Studie

Ziel dieser Arbeit ist es, den genauen Einfluss von lexikalischer und grammatikalischer Verarbeitung exemplarisch anhand der Verarbeitung des englischen Pluralmorphems -s zu erfassen. Das Pluralmorphem -*s* kommt im Englischen in den meisten Fällen vor, wie z. B. in dem Wort “cats”. In dieser Studie wird gezeigt, dass das Pluralmorphem nicht ausschließlich grammatikalisch verarbeitet wird. Die Sprachverarbeitung bezieht sich in meinen Experimenten auf die Sprachperzeption beim Lesen von geschriebenen Wörtern und Wortkombinationen, gemessen anhand von Reaktionszeiten.

Die Einführung einer neuen Struktur im Sprachunterricht erfolgt oftmals durch vermehrtes Üben einer neuen Regelanwendung. Grammatik wird als eine strukturelle Eigenschaft gesehen und daher funktional und auch prozedural bei der Verarbeitung von den inahltlichen Eigenschaften getrennt. In der besprochenen Literatur werden viele Gründe genannt, eine solche Trennung vorzunehmen. Auch in Sprachverarbeitungsstudien gibt es Evidenz für eine Trennung zwischen lexikalischer und grammatikalischer Verarbeitung, sowohl bei der Lokalisation im Gehirn, als auch temporal im Zeitablauf der Verarbeitung. Diese Trennung zwischen Inhalt und Form ist oftmals eine Vereinfachung der tatsächlichen Abläufe. In dieser Arbeit möchte ich zeigen, dass die Verarbeitung des einfachen Pluralmorphems "s" nicht so leicht in eine rein grammatikalische Verarbeitung zu kategorisieren ist. Das Pluralmorphem hat durchaus auch semantische Eigenschaften und ist in den Reaktionszeitenexperimenten nicht eindeutig temporal von inhaltlicher Verarbeitung zu trennen. Vielmehr spielen eine Vielzahl von Faktoren eine Rolle, die ich in den Experimenten identifizieren konnte.

Die Grundhypothese dieser Studie ist, dass Pluralverarbeitung des englischen Pluralmorphems "s" keine rein grammatikalische Angelegenheit ist und sich dies auch experimentell zeigen lässt. In Kapitel 5 werden überblicksartig neueste Erkenntnisse aus der Literatur dargestellt, die eine strikte Trennung zwischen grammatikalischer und lexikalischer Verarbeitung problematisch erscheinen lassen. Die ausgewertete Literatur, die in Kapitel 5 präsentiert wird, zeigt, dass die Verarbeitung von Flexionsmorphemen sehr wohl durch gewisse semantische oder auch formale Faktoren (z. B. der Unterschied zwischen regelmäßigem bzw. unregelmäßigem Flexionsmorphem, typologische Unterschiede bei den Flexionsmorphemen, Frequenzen) beeinflusst wird. Diese Erkenntnis wiederum wirft die Frage auf, welche Faktoren eine eher lexikalische bzw. eine eher grammatikalische Verarbeitung begünstigen. Die hier vorliegende Studie kann diese Frage ansatzweise beantworten und stellt fest, dass Verarbeitungszeiten von diversen Faktoren, wie z.B. Frequenz, Dominanz und konzeptuelle Assoziation, abhängig sind.

Eine weitere Frage, die in dieser Arbeit beantwortet wird, betrifft die Verarbeitung von Flexionsmorphemen bei Sprachlernern. Wenn es einen Verarbeitungsunterschied zwischen Muttersprachlern und Lernern gibt, dann ist es wichtig zu wissen, wie die oben genannten Faktoren (Frequenz, Dominanz, konzeptuelle Assoziation) den Verarbeitungsweg bei den jeweiligen Gruppen beeinflussen. Sollte der Einfluss dieser Faktoren bei steigender Sprachfähigkeit von Lernern des Englischen stetig wachsen, dann kann man durch Verarbeitungszeiten eine gewisse Art der Sprachkompetenz messen. Sollten diese Faktoren nicht auftreten, obwohl sie bei der Sprachverarbeitung von Muttersprachlern nachgewiesen sind, diese Lerner jedoch die Pluralformen korrekt benutzen können, so kann man einen qualitativen Verarbeitungsunterschied zwischen

Muttersprachlern und Lernern feststellen. Theoretisch wäre es möglich, dass man nur quantitative Unterschiede zwischen den verschiedenen Lernerstufen erkennen könnte, dann aber bei den Muttersprachlern eindeutig qualitative Verarbeitungsunterschiede sichtbar wären. Möglich ist aber auch, dass die Unterschiede zwischen Lernern und Muttersprachlern rein quantitativ sind, aber beispielsweise die Verarbeitungszeiten bei beiden Gruppen stetig schneller werden. Die Ergebnisse der beiden Experimente in dieser Studie sowohl für Lerner als auch für Muttersprachler deuten darauf hin, dass die Annahme von zwei strikt getrennten Prozessen im Sinne von lexikalischer und grammatikalischer Verarbeitung zu bezweifeln ist. Die Ergebnisse zeigen auch, dass die Annahme von einer Zweiteilung zwischen Muttersprachlern und Lernern bei der Sprachverarbeitung nicht belegt werden kann. Die Ergebnisse zeigen, dass alle Gruppen ein eigenes Verhalten vorweisen, was auch für die individuellen Probanden angenommen werden muss.

Ergebnisse aus der Literatur

Wie im ersten Kapitel bei der Auswertung der aktuellen wissenschaftlichen Beiträge diskutiert wird, werden grammatikalische- und lexikalische Verarbeitung weitgehend als zwei voneinander relativ unabhängige Prozesse betrachtet. Dieses gilt sowohl für die Verarbeitungszeit (gemessen anhand von Reaktionszeiten bei online Sprachverarbeitung) als für den Verarbeitungsort (gemessen durch Magnetresonanzverfahren oder auch durch elektrophysiologische Respons auf sprachliche Stimuli). Die Literatur bietet reichhaltige Beispiele, in denen von einer funktionalen Trennung zwischen Grammatik und Lexikon ausgegangen wird. Die Trennung in der Verarbeitung (lokal oder temporal) zwischen Grammatik und Lexikon spiegelt sich beispielsweise in der Unterscheidung zwischen Flexionsmorphemen und derivationalen Morphemen (Perlmutter, 1988; Badecker & Caramazza, 1989; Schriefers, Friederici, und Graetz, 1992), zwischen inherenten und kontextuellen Morphemen (Booij, 1993; Pienemann, 1998; Tobin, 2000; Hirtle, 2000; Hay, 2001) oder in der Unterscheidung zwischen regelmäßigen und unregelmäßigen Flexionsmorphemen (Sonnenstuhl, Eisenbeiss & Clahsen, 1999; Lück, Hahne & Clahsen, 2001; Ullman, 2004) wieder. Diese Studien zeigen, dass allein die Verarbeitung des Flexionsmorphems, das lange Zeit als rein grammatikalisch galt, in sich schon nicht einheitlich verläuft. Vielmehr wird das Flexionsmorphem auf verschiedene Arten verarbeitet.

Diese Diversität gilt auch für typologische Unterschiede: Flexionsmorpheme verhalten sich unterschiedlich in verschiedenen Sprachen. In Kapitel 5 wurden einige Beispiele aus dem Finnischen und dem Italienischen gezeigt (siehe 5.2.2.), die belegen, dass die Verarbeitung dieser Morpheme von Faktoren wie Frequenz oder Produktivität abhängt, nicht jedoch von der Kategorie der Flexionsmorpheme oder Derivationsmorpheme an sich. Obwohl es Hinweise für eine pluriforme Verarbeitung gibt, die nicht strikt grammatikalisch oder lexikalisch verläuft, so gibt es andererseits auch Studien, die eine funktionale Trennung der beiden belegen. Eine solche Trennung wurde auch in einer Neuroimaging-Studie von Ullman (2004) gefunden. Ullman geht davon aus, dass es zwei Gedächtnissysteme im Gehirn gibt, die funktional getrennt sind und sich im Großen und Ganzen in der Unterscheidung zwischen Auswendiglernen und Regelanwendung widerspiegeln. Für englische Pluralformen würde dies bedeuten, dass ein Nomen im Singular im Lexikon gespeichert wird, die Pluralform jedoch aus diesem Nomen im Lexikon und einer Regelanwendung für das Pluralmorphem gebildet wird. Diese Regelanwendung wäre dann eine strikt grammatikalische Verarbeitung, die zeitlich nach dem lexikalischen Zugriff auf das Nomen im Lexikon geschieht.

Die Ergebnisse meiner Arbeit legen die Vermutung nahe, dass es eine gewisse Art von Überlappung zwischen lexikalischer und grammatikalischer Verarbeitung gibt, die abhängig von den Verarbeitungsumständen ist. Die beiden Verarbeitungsmodule sind nicht streng trennbar und auf keinen Fall exklusiv verantwortlich für bestimmte Verarbeitungsbereiche. Ein und dasselbe Morphem kann je nach Situation auf entweder lexikalische oder grammatikalische Art verarbeitet werden. Ähnlich wie in der Diskussion um die Rolle des Erst- und des Zweitsprachlexikons bei der Bestimmung der Zweisprachigkeit bei Weinreich (1953) und später auch aufgegriffen durch Forscher wie de Groot (1993) und Kroll (1993) (siehe 2.3.2 ff), bei der die Problematik einer eindeutigen Einteilung in die verschiedenen Arten des Bilingualismus hervorgehoben wird, plädiere ich dafür, dass die richtige Mischung aus grammatikalischer bzw. lexikalischer Verarbeitung nur zu bestimmen ist, wenn man Eigenarten und Sprachstufe des individuellen Lerners berücksichtigt. Die Ergebnisse aus meinen beiden Experimenten belegen genau diese Annahme: Verarbeitungsunterschiede hängen von diversen formalen und konzeptuellen Faktoren ab.

Die Arbeiten von Sereno und Jongman (1997) und von Hay (2001) kommen zu dem Schluss, dass eine strikte Zweiteilung bei der morphologischen Verarbeitung nicht möglich ist. Auch Ullman (2004) erwähnt, dass die beiden zu trennenden Gedächtnissysteme, die mit lexikalischer und grammatikalischer Verarbeitung korrelieren, nicht sprachspezifisch und daher auch nicht strikt trennbar in ihrer Domänenspezifität sind (siehe 2.3). Jones (2004) plädiert für ein Gedächtnissystem, das in ein semantisches und episodisches Gedächtnis getrennt ist, unabhängig davon, ob etwas sprachlich verarbeitet wird oder nicht. Die genannten Beispiele aus der Literatur zeigen, dass die beiden linguistischen Verarbeitungsmodule organisch, und funktional unterschiedlich arbeiten, ihnen jedoch immer eine gewisse Art von Überlappung zugestanden werden muss. Auch die englische Pluralbildung kann daher kein rein grammatikalischer Prozess sein; sie muss von unterschiedlichen Faktoren abhängig sein, die die Pluralbildung eher lexikalisch bez. eher grammatikalisch verlaufen lassen.

Überblick über die einzelnen Kapitel

Nach Erläuterung der Problematik und der Forschungsfragen werden im folgenden die Inhalte der Kapitel zusammengefasst bevor die Ergebnisse präsentiert werden.

In der Einführung stelle ich die oben beschriebene Problematik dar, auf der meine Forschungsfragen beruhen. Das Hauptanliegen meiner Arbeit ist der Beweis der These, dass das englische Pluralmorphem nicht den strikt grammatikalischen Verarbeitungsweg durchläuft, sondern dass eine ganze Reihe von inhärenten und externen Faktoren die Art der Sprachverarbeitung beeinflussen. Des Weiteren möchte ich im Rahmen meiner Arbeit die Art dieser Faktoren näher bestimmen und dabei auch klären, ob diese Faktoren bei Muttersprachlern und Sprachlernern gleichermaßen eine Rolle spielen.

Um die oben genannten Forschungsfragen zu klären, werden in der vorliegenden Arbeit flektierte Pluralformen getestet, während Probanden einzelne Wörter online verarbeiten und sich der eigentlichen Absicht des Experiments nicht bewusst sind. Die Rolle der Flektionsmorpheme bei der Sprachverarbeitung im Lexikon und beim lexikalischen Zugriff müssen daher genau erklärt werden. Aus diesem Grund beschreibt Kapitel 2 ein psycholinguistisches Modell der lexikalischen Organisation. In diesem Modell wird deutlich, in welchem Verhältnis Grammatik und Lexikon zueinander stehen und warum Flexionsmorpheme als eine eigene Klasse - im Gegensatz zu den Derivationsmorphemen - definiert sind. Im zweiten Kapitel wird die Literatur besprochen, die zu der Trennung zwischen Lexikon und Grammatik kritisch Stellung nimmt. Kapitel 3 konzentriert sich auf die Rolle des Lexikons und diskutiert einige der bekannten Theorien zum mentalen lexikalischen Zugriff. Diese lexikalische Verarbeitung geschieht innerhalb des großen mentalen Organisationsmodells, das bereits in Kapitel 2 dargestellt wird. Obwohl es problematisch erscheint die Rolle des Lexikons und seine Funktion im Gesamtmodell von den lexikalischen Zugriffstheorien zu trennen, so schließen die meisten dieser Theorien die größere Gesamtorganisation nicht mit ein. Daher spiegelt sich diese Trennung von Organisation und Funktion auch in den beiden Kapiteln wieder.

Im vierten Kapitel gebe ich eine kurze Übersicht über die Bedeutung des Pluralmorphems in der englischen Sprache. Dieses Kapitel beschreibt die Pluralanwendung in gesprochener Sprache und die Bedeutung des Numerus auch im Vergleich zu anderen Sprachen oder Dialekten. Hierbei wird deutlich, dass es sich beim Plural nicht lediglich um eine grammatikalische Form handeln kann, denn er hat durchaus lexikalische Aussagekraft. So gibt es Sprachen, die den Plural nicht morphosyntaktisch kodieren, andere Sprachen unterscheiden neben Singular und Plural auch den Dual. Die Funktion des Plurals ist also keine rein grammatikalische, und in manchen Fällen kann der Einsatz des Pluralmorphems anstelle eines Zahlwortes als Pluralmarker einen semantischen Unterschied ausmachen. Diese Übersicht über den Gebrauch des Plurals stützt meine These, dass Sprachverarbeitung nicht rein mathematisch nach Regelanwendung und Lexikon getrennt werden kann, sondern dass auch konzeptuelle Faktoren eine Rolle spielen müssen.

Kapitel 5 kehrt zu den formalen Eigenschaften der Pluralverarbeitung zurück und es werden einige Studien beschrieben, die sich mit der Trennung zwischen lexikalischer und grammatikalischer Pluralverarbeitung befassen und schwerpunktmäßig die Verarbeitung des regelmäßigen Plurals untersuchen. Die diversen experimentellen Studien zeigen, dass es einen graduellen Unterschied zwischen Grammatik und Lexikon geben muss. Dieses Kapitel bildet zugleich den theoretischen Hintergrund für die Entwicklung meiner eigenen Experimente. In Kapitel 6 wird der Zweck, der Versuchsaufbau und die genauen Abwägungen im Testdesign der beiden Experimente erklärt, die für diese Studie benutzt wurden: ein Wortentscheidungsexperiment und ein Grammatikalitätsexperiment. Kapitel 7 fasst die Ergebnisse dieser beiden Experimente zusammen. In Kapitel 8 folgt dann eine allgemeine Diskussion dieser Ergebnisse im Hinblick auf die bereits beschriebenen Theorien und die aufgestellten Hypothesen und es werden auch die Grenzen dieser Untersuchung beschrieben, die neue Forschungsfragen aufrufen. Es folgen einige Anregungen für weitere Untersuchungen, die auf den bisherigen Ergebnissen aufbauen könnten.

Beschreibung des Wortentscheidungsexperiments

Im folgenden wird das erste meiner Experimente, das Wortentscheidungsexperiment, kurz zusammengefasst. Bei diesem Experiment sitzt der Proband am Bildschirm und bekommt nacheinander unterschiedliche Wörter und Pseudowörter zu sehen. Jedes mal soll der Proband entscheiden, ob das gelesene Wort ein real existierendes Wort ist oder nicht. Dabei registriert der Computer die Reaktionszeit des Wortzugriffs in Millisekunden. Der Proband weiß aber nicht, dass es bei der Forschungsfrage nicht um den Unterschied zwischen Wort und Pseudowort geht, sondern dass sich vielmehr unter den reellen Wörtern einige Stimuluswörter befinden, die aufgrund bestimmter Kriterien ausgesucht wurden. Von den rund 400 Wörtern, die der Proband zu sehen bekommt, sind nur 70 Wörter eigentliche Stimuluswörter, die für meine Pluraluntersuchung relevant sind. Warum sich ein Wortentscheidungsexperiment mit einem solchen Aufbau für meine Untersuchung sehr gut eignet und welche Effekte bei den Reaktionszeiten normalerweise auftreten wird ausführlich in Kapitel 6 beschrieben. Die Testbedingungen wurden in einer Pilotstudie getestet und daraufhin modifiziert (Instruktionen und Timing am Ende des Feedbacks). Zwei Probandengruppen nahmen an dem Test teil: englische Muttersprachler sowie Lerner des Englischen. Dabei gab es verschiedene Untergruppen: Es gab 29 Muttersprachler (Schüler) aus der Altersgruppe der 16 bis 17-jährigen. 34 Muttersprachler (Soldaten) waren zwischen 18 und 39 Jahren alt. Die Englischlerner waren in Schulklassen unterteilt: durchschnittlich etwa 20 Schüler je Schulklasse. Die Klassen 7 bis 12 wurden getestet (genaue Angaben zur Methodik in Kapitel 6).

Die Stimuluswörter wurden aufgrund der Frequenz ausgesucht, da Frequenz ein Faktor ist, der bei diesem Experiment untersucht werden soll. Um die Frequenz der jeweiligen Wörter zu bestimmen, wurde der CELEX Datenkorpus benutzt, in dem 160000 Wortformen für die englische Sprache mit Angabe der jeweiligen Frequenz enthalten sind.

Hier wird sowohl die absolute als auch die logarithmische Frequenz angegeben. Aufgrund der menschlichen Perzeption wird hier bei der Angabe der drei unterschiedlichen Frequenzstufen von der logarithmischen Frequenz ausgegangen. Im Kapitel 6 wird detailliert beschrieben, warum diese Frequenzen für diese Studie relevant sind.

Neben der Frequenz ist auch die Dominanz ein entscheidender zu testender Faktor in meinen Experimenten. Ein Wort kann singular-dominant oder plural-dominant sein, je nach dem in welcher der beiden Formen es häufiger auftritt (nach dem CELEX Datenkorpus). Das ist der so genannte Dominanzfaktor in meinem Experiment. Um diesen zu untersuchen, mussten die Wörter auf ihre Dominanz hin ausgewählt werden. Dabei konnten die Stimuluswörter in acht verschiedene Dominanzverhältnisse aufgeteilt werden mit Berücksichtigung der jeweiligen Frequenz von Wortstamm und Pluralform. Für jede Wortstammgruppe (also hohe Wortstammfrequenz, mittlere Wortstammfrequenz und niedrige Wortstammfrequenz) gab es jeweils drei verschiedene Pluralfrequenzen (hoch-, mittel-, und niedrig-frequente Pluralform). Bei der Gruppe mit niedriger Wortstammfrequenz gab es keine Einträge für hohe Pluralfrequenz, weshalb es auch insgesamt keine neun Gruppen, sondern acht Gruppen mit unterschiedlichen Dominanzverhältnissen gibt.

Nun kann es innerhalb der gleichen logarithmischen Frequenzebene (z. B. hoch-frequent) bei Wortstamm und Pluralform zu großen Unterschieden kommen. Manchmal liegen die Frequenzen an den äußersten Enden einer Frequenzebene (z. B. Wortstamm unteres Ende der hoch-frequenten Ebene und Pluralform an dem oberen Ende der hoch-frequenten Ebene). Es kann aber auch sein, dass innerhalb einer Ebene beide Formen innerhalb des gleichen Bereichs liegen. Da aus der Literatur (siehe Kapitel 6 für nähere Erklärungen) hervorgeht, dass die relative Frequenz bei der Verarbeitung eine Rolle spielt, musste auch hier eine Lösung gefunden werden. Alle Wörter, deren Stamm und Pluralform sich in unterschiedlichen Ebenen befinden sind jeweils dominant in eine Richtung mit hoher Relativität. Alle anderen Wörter wurden nochmals genauer betrachtet und in singular- bzw. plural-dominante Gruppen mit niedriger Relativität unterteilt. Die verbliebenen Gruppen sind folgende:

Table 14.

Gruppe	Stamm frequenz	Dominanz	relative Dominanz
1	hoch	plural	niedrig
2	hoch	singular	hoch
3	mittel	singular	niedrig
4	mittel	singular	niedirg
5	mittel	plural	niedrig
6	mittel	singular	hoch
7	niedrig	plural	hoch
8	mittel	singular	niedrig
9	niedrig	plural	niedrig
10	niedrig	singular	niedrig

Für jede Gruppe wurden sechs bis acht Stimuluswörter ausgesucht, insgesamt 70 Wörter. Den zwei Versionen des Experiments wurden jeweils deren Singular- und Pluralformen zugeteilt (also insgesamt 140 Wortformen), beliebig verteilt auf die beiden Testversionen. Wenn eine Singular oder Pluralform in Version A vorkommt, so kommt das entsprechende Singular- oder Pluraläquivalent in Version B vor. Niemals kommen Singular und Pluralform des gleichen Wortes in einer Testversion vor. Jede Version hat außerdem 120 existierende Füllwörter sowie 190 Pseudowörter. Der Test besteht insgesamt aus 380 Wörtern, auf die die Probanden reagieren soll. Vor jedem Experiment erscheinen zusätzlich noch 10 Übungswörter, damit sich der Proband an den Ablauf gewöhnen kann. Diese gehen nicht mit in die Auswertung ein.

Beschreibung des Grammatikalitätsexperiments

Sinn dieses zweiten Experiments ist es, nach den probabilistischen Faktoren wie Frequenz und Dominanz auch die Wortbedeutungsebene zu berücksichtigen. Es soll untersucht werden, ob konzeptuelle Faktoren die Verarbeitung des Pluralmorphems beeinflussen. Bei der Pluralverarbeitung ist es das Konzept der Numerosität. Diese wird ausführlich in

Kapitel 4 diskutiert. In dem Grammatikalitätsexperiment geht es darum, konzeptuelle Duale mit konzeptuellen Pluralen zu vergleichen. Ein konzeptueller Dual ist eine Pluralform, bei der man sich automatisch eine Zweiheit bzw. ein Paar vorstellt, wie z.B. "eyes" (Augen) oder "wings" (Flügel). Mit konzeptuellen Pluralen sind Pluralformen gemeint, die im Allgemeinen als Vielzahl auftreten, wie z. B. "tears" (Tränen). "miles" (Meilen) oder "words" (Wörter). Ob diese Information einen Einfluss auf die Verarbeitungszeit hat, soll in dem grammatikalischen Beurteilungsexperiment getestet werden.

Das Experiment verläuft wie das Wortentscheidungsexperiment. Der Proband sitzt am Bildschirm und reagiert per Knopfdruck auf nacheinander erscheinende Stimuluswortpaare, meist Nominalphrasen. Diese setzen sich aus einem Zahlwort und einem Nomen in der Pluralform zusammen. Auch die Füller und Pseudopaare sind meist Nominalphrasen, manchmal sind die Distraktoren auch Verbalphrasen. Bei diesem Experiment entscheiden die Probanden nicht, ob es sich um existierende Wörter handelt, sondern ob die Wortpaare (z. B. Zahlwort und Nomen: "three fingers") grammatikalisch zusammenpassen. Ein Paar wie "three finger" wäre also ein Paar, das grammatikalisch nicht zusammenpasst. Der Computer registriert die Zeit in Millisekunden, die der Proband benötigt, um seine grammatikalische Entscheidung für die jeweilige Phrase zu treffen. Für die Auswahl der entscheidenden Stimulusnominalphrasen, die in die Auswertung eingehen, habe ich Pluralformen in Kombination mit Zahlwörtern verwendet. Dabei habe ich konzeptuelle Duale und konzeptuelle Plurale mit zwei Typen von Zahlwörtern benutzt. Die beiden Zahlworttypen sind zum einen ein duales Zahlwort (also "zwei") und zum anderen Pluralzahlwörter (mehr als zwei), alle waren einsilbig. Aus diesen Kombinationen entstanden vier verschiedene Stimulustypen:

- Dualzahlwort + konzeptueller Dual (z. B. *two eyes*)

- Dualzahlwort + konzeptueller Plural (z. B. *two tears*)

- Pluralzahlwort + konzeptueller Dual (z. B. *ten eyes*)

- Pluralzahlwort + konzeptueller Plural (z. B. *ten tears*)

Für die Dualnomen erwarte ich schnellere Reaktionszeiten bei den Nominalphrasen mit Dualzahlwort als bei den Nominalphrasen mit Pluralzahlwörtern. Bei den konzeptuellen Pluralen vermute ich, dass es keinen Unterschied gibt, da es beim konzeptuellen Plural auf die genaue Anzahl oft nicht ankommt und außerdem geringe Anzahlen, wie "drei" oder "fünf" genauso unnatürlich sind in Kombination mit z. B. "Tränen", wie das Zahlwort "zwei". Die Verarbeitungszeit würde dann also nicht vom Zahlwort abhängen, sondern von der Numerosität des Nomens (ob konzeptueller Dual oder konzeptueller Plural). Die Probanden wissen nicht um die eigentliche Absicht dieses Tests sondern konzentrieren sich auf den grammatikalischen Aspekt.

Die Pluralformen der konzeptuellen Duale und der konzeptuellen Plurale wurden sorgfältig nach Numerosität und Frequenz ausgewählt. Auch die Dominanz der Wörter wurde kontrolliert, damit eventuelle Ergebnisse keine Artefakte sind, die aufgrund der Dominanz der Wörter auftauchen. So ist die Hälfte aller Stimuluswörter der konzeptuellen Duale singular-dominant und die andere Hälfte ist plural-dominant. Die konzeptuellen Plurale sind per Definition plural-dominant. Jede Testversion beinhaltet 13 konzeptuelle Dualwörter und 12 konzeptuelle Pluralwörter mit entweder einem dualen oder einem Pluralzahlwort (zufällig verteilt auf die beiden Testversionen). Die Testversionen unterscheiden sich also lediglich zwischen den Zahlwörtern der Stimulusnominalphrasen. Die restlichen Wortpaare der insgesamt 195 Stimulusphrasen sind gleich.

Die Probanden sind die gleichen Gruppen wie beim bereits beschriebenen Wortentscheidungsexperiment, bis auf zwei Lernerklassen (eine Klasse 8, eine Klasse 9), die beim zweiten Experiment durch das Schreiben von Klassenarbeiten verhindert waren. Insgesamt nahmen 169 Probanden an diesem Experiment teil: 63 Muttersprachler und 106 Lerner der Klassen 7, 9, 10, 11 und 12.

Ergebnisse beim Wortentscheidungsexperiment

Allgemeine Reaktionszeiten

Bei den allgemeinen Reaktionszeiten zeigt sich, dass die britischen Schüler am schnellsten auf die Stimuli reagieren. Ihre Reaktionszeiten sind bedeutend schneller als die aller Lernergruppen und vor allem auch als die der britischen Soldaten. Die Soldaten werden in den langsamen Reaktionszeiten nur von den Siebtklässlern (Lerner) übertroffen. Bei den Lernern gibt es keinen Zusammenhang zwischen Klassenstufe und stetig abnehmender Reaktionszeit. Man würde erwarten, dass alle Muttersprachler schneller reagieren als die Lerner. Die Reaktionszeiten der Soldaten widersprechen dieser Annahme jedoch. Anzunehmen ist, dass sie aufgrund ihrer alltäglichen Aufgaben nicht gewohnt sind, auf schriftliche Anweisungen am Bildschirm zu reagieren. Man darf aus dieser Untersuchung daher auch nicht schließen, dass sich die Automatisierung von Sprache in den Reaktionszeiten widerspiegelt. Die sprachliche Automatisierung der Soldaten wird in diesen Ergebnissen nicht reflektiert, Reaktionszeiten sind daher sehr aufgabenabhängig. Es ist daher fraglich, ob Experimente mit Reaktionszeiten ein geeignetes Instrument sind, um etwas über den Verarbeitungsunterschied zwischen Muttersprachlern und Sprachlernern auszusagen.

Richtige Wörter und Unsinnwörter

Alle Gruppen sind sehr konsequent bei den Reaktionszeiten auf richtige Wörtern und Pseudowörter. Obwohl die Pseudowörter nicht in die Auswertung mit eingehen, so ist es interessant, zu sehen, ob die Probanden sich hier wie in der Literatur verhalten und der Test somit valide und verlässlich ist. Richtige Wörter werden schneller verarbeitet als Pseudowörter und damit bestätigt dieses Experiment alle Vorhersagen aus der Literatur und kann als verlässlich gesehen werden.

Numeruseffekt

Keine der Muttersprachlergruppen zeigt einen Numerus-Effekt, d. h. weder die britischen Schüler noch die Soldaten verarbeiten Singularformen signifikant schneller als Pluralformen. Man kann davon ausgehen, dass die Muttersprachler so vertraut mit beiden Wortformen sind, dass es bei der Verarbeitung nicht zu signifikanten Unterschieden kommt. Bei den Lernern zeigt sich der Effekt nur bei den Probanden der Klassenstufe 7 und 11. Sie verarbeiten die Singularformen um 20 ms schneller als die Pluralformen. Bei den Lernern aus Klasse 7 könnte man mutmaßen, dass sie im Sprachinput häufiger den Singular- als den Pluralformen ausgesetzt wurden, und diese daher nicht so automatisiert sind. Bei den

Lernern aus Klasse 11 ist dieser Grund jedoch eher unwahrscheinlich, sonst würde das gleiche Phänomen auch bei den anderen darunter liegenden Klassen auftauchen. Diese Ergebnisse lassen sich daher nicht so einfach erklären.

Interaktion zwischen Numerus und Dominanz

Die Interaktion zwischen Numerus und Dominanz ist sehr stark bei den britischen Schülern ausgeprägt, die einen großen Unterschied bei den plural-dominanten Nomen zeigen, d.h. ziemlich niedrige Reaktionszeiten für plural-dominante Nomen in der Pluralform und vergleichsweise hohe Reaktionszeiten für plural-dominante Nomen in der Singularform. Es ist weiterhin ein starker Effekt für die singular-dominanten Wörter zu beobachten, wobei dieser vor allem bei den Pluralformen eine Rolle spielte.

Die britischen Soldaten zeigen keinen Interaktionseffekt. Bei sowohl singular- als auch plural-dominanten Wörtern haben sie die Singularformen schneller verarbeitet. Aber auch hier ist der Unterschied für die Pluralformen größer als für die Singularformen, d.h. die Rekationszeiten zwischen pluraldominanten und singulardominanten Pluralformen unterscheiden sich stärker als die Reaktionszeiten zwischen singulardominanten und pluraldominanten SIngularformen.

Die meisten Lernergruppen zeigen keinen Interaktionseffekt. Die singular-dominanten Wörter werden aber im Singular schneller verarbeitet als in der Pluralform. Die einzige Lernergruppe mit einem Interaktionseffekt ist Klasse 9. Hier ist der Effekt ebenfalls stärker für die Pluralformen als für die Singularformen.

Zusammenfassend lässt sich erstens feststellen, dass es bei der Interaktion zwischen Numerus und Dominanz einen qualitativen Unterschied zwischen den Sprachlernern und den Muttersprachlern allein in der Existenz dieses Interaktionseffekts gibt, der bei den Lernern meist nicht vorhanden ist, bei den britischen Schülern jedoch wohl. Zweitens kann man einen Unterschied feststellen in der Art dieses Effekts. Die britischen Schüler zeigen einen starken Effekt bei den plural-dominanten Wörtern, die Lerner zeigen starke Effekte bei singular-dominanten Wörtern. Diese Diskrepanz lässt sich dadurch erklären, dass die Lerner die singular-dominanten Wörter im Singular weit genug automatisiert haben und, dass deren Pluralformen noch eine größere Verarbeitungshürde darstellen, die sich in den

Reaktionszeiten widerspiegelt. Dieser Effekt ist bei den plural-dominanten Wörtern nicht festzustellen, da die plural-dominanten Pluralformen nicht in der gleichen Weise automatisiert werden wie die singular-dominanten Singularformen, so dass es keinen Verarbeitungsvorteil für plural-dominante Pluralformen im Vergleich zu deren Singularform gibt. Dieser Effekt tritt bei den britischen Schülern jedoch sehr wohl auf. Bei ihnen wird die jeweilig häufigere Form (also singular-dominante Singulare und plural-dominante Plurale) schneller verarbeitet.

Die Soldaten zeigen jedoch keinen Interaktionseffekt sondern nur einen Unterschied zwischen singular-dominanten Singularen und singular-dominanten Pluralen. Dieses Ergebnis ist vergleichbar mit den Lernergruppen, die Soldaten zeigen hier auch das zu erwartende Ergebnis. Ein Ausbleiben eines Interaktionseffekts könnte man auch hier darauf zurückführen, dass Soldaten meistens keine geschriebenen Wörter in ihrer alltäglichen Routine verarbeiten. Der große Verarbeitungsunterschied zwischen den englischen Muttersprachlern ist weiterhin interessant. Auch dabei stellt sich die Frage, ob sich anhand solcher Experimente überhaupt eine Aussage über Muttersprachler im Vergleich zu Lernern treffen lässt. Man sollte jede Sprachverarbeitung eher als ein großes Kontinuum sehen, in dem jeder Sprecher oder auch jede Gruppe seine eigenen Koordinaten finden muss. Die Tatsache, dass bei der Interaktion wieder eine Lernergruppe aus der mittleren Klassenstufe in den Ergebnissen auffällt, zeigt, dass sich die Lernergruppen auch nicht kontinuierlich entwickeln.

Interaktion Numerus und relative Dominanz

Die Interaktion zwischen Numerus und relativer Dominanz bezieht sich auf eine Untergruppe aus den originalen Zielwörtern, nämlich die, der hochdominanten (sowohl singular- als auch pluraldominant). Dieser Interaktionseffekt sollte daher auch stärker sein als der vorherige zwischen Numerus und absoluter Dominanz.

Tatsächlich verhalten sich die britischen Schüler hier wie erwartet, vor allem bei den Pluralnomen, bei denen die plural-dominanten Nomen viel schneller verarbeitet werden als die singular-dominanten Nomen. Bei den britischen Soldaten kann der Interaktionseffekt in diesem Fall auch festgestellt werden. Bei den Soldaten ist der Effekt bei den singular-dominanten Wörtern jedoch größer. Wichtig ist dabei, dass ein signifikanter

Interaktionseffekt nachgewiesen werden kann. Keine der Lernergruppen zeigt einen Interaktionseffekt. Man kann aber feststellen, dass die Unterschiede bei den singular-dominanten Wörtern größer sind als bei den plural-dominanten Wörtern.

Dadurch, dass es sich um eine Untergruppe der vorherigen Zielwörter handelt, ist es erfreulich, dass sich klare Verhältnisse herauskristallisieren, die den vorherigen Interaktionseffekt verstärken bzw. - bei den Lernern - schwächen. Dominanz beruht auf probabilistischen Faktoren. Sobald sich das Maß der Dominanz verstärkt, zeigt sich das in den Reaktionszeiten. Probabilistische Faktoren sind daher ein ernstzunehmender Faktor bei der Sprachverarbeitung. Dabei ist es irrelevant, dass der Unterschied zwischen absoluter und relativer Dominanz bei den Sprachlernern keine Rolle spielt. Diese sind noch nicht so anfällig für probabilistische Faktoren, da die Stimuluswörter in diesem Experiment sowieso aus einem Muttersprachlerkorpus stammen (CELEX Database). Bei den Lernern ist jedes Vokabular verhältnismäßig niedrig-frequent daher nicht in der Form automatisiert wie bei den Muttersprachlern. Plural-dominante Nomen sind nicht automatisch auch im Lernerinput pluraldominant, oft werden neue Vokabeln in der Singularform gelernt und gespeichert. Die probabilistischen Faktoren können daher nicht ohne weiteres von Muttersprachler auf Lerner übertragen werden; Lerner können eine ganz eigene probabilistische Denkweise besitzen.

Frequenzeffekt

Es werden signifikante Frequenzeffekte für alle Frequenzebenen (hoch, mittel und niedrig-frequent) bei den britischen Schülern gefunden, d. h. je frequenter die Wörter, desto schneller werden sie verarbeitet. Bei den britischen Soldaten wird ein allgemeiner Frequenzeffekt festgestellt, der allerdings näher betrachtet nur zwischen hoch- und niedrig-frequenten bzw. mittel- und niedrig-frequenten Wörtern signifikant ist. Für die Soldaten tauchen mittel-frequente Wörter im Sprachgebrauch häufig genug auf, um keinen großen Effekt im Vergleich zu den hoch-frequenten Wörtern zu verursachen und daher auf gleiche Weise verarbeitet zu werden.

Die Lerner zeigen alle einen signifikanten Frequenzeffekt, wenn auch nicht zwischen allen Ebenen. Alle Lerner verarbeiten hoch-frequente Wörter viel schneller als mittel- und niedrig-frequente Wörter. Dieser Effekt wird durch vorherige Experimente aus der Frequenzforschung auch erwartet. Die Lerner unterscheiden jedoch wohl zwischen hoch- und mittel-frequenten Wörtern. Nur die hoch-frequenten Wörter sind im Sprachgebrauch auch wirklich häufig genug, um bei den Reaktionszeiten Effekte zu erzielen. Dies ist wiederum gut zu erklären, da die Frequenzen auf Muttersprachlerfrequenzen (CELEX Database) beruhen und vom normalen Sprachgebrauch von Muttersprachlern ausgehen.

Ergebnisse beim Grammatikalitätsexperiment

Konzeptuelle Effekte bei der Sprachverarbeitung

Bei fast allen Lernern und bei allen Muttersprachlern gibt es keinen signifikanten Unterschied in der Verarbeitungszeit zwischen den konzeptuellen Pluralen mit dualem Zahlwort und den konzeptuellen Pluralen mit Pluralzahlwort (z.B. *two teeth* versus *ten teeth*). Nur die Lerner der Klasse 11 verarbeiten die Plurale mit dem dualen Zahlwort signifikant schneller. Obwohl es bei den meisten Gruppen keine signifikanten Effekte gibt, so scheinen die Gruppen im allgemeinen das duale Zahlwort zu favorisieren, da die Verarbeitungszeiten durchgehend schneller beim dualen Zahlwort sind, wenn auch nicht signifikant. Dieses Phänomen kann durch den sogenannten *Subitising-Effekt* (Dehaene, 2001) erklärt werden. Demnach können Menschen alle Größen unter vier sofort identifizieren. Dieser Effekt beschränkt sich nicht auf die Anzahl von Objekten sondern gilt auch für deren symbolische Präsentation durch Zahlen oder Zahlwörtern. Alle Größen unter vier haben also einen gewissen Verarbeitungsvorteil, was erklären kann, warum das Zahlwort “two” in Kombination mit konzeptuellen Pluralen im Schnitt schneller verarbeitet wird als die anderen getesteten Zahlwörter unter zehn. Allerdings ist dieser Effekt scheinbar nicht groß genug, um signifikante Unterschiede zu bewirken (außer bei Klasse 11).

Bei den konzeptuellen Dualen sind die Ergebnisse interessanter. Beide Gruppen der Muttersprachler sind signifikant schneller bei der Reaktion auf ein duales Zahlwort in Kombination mit dualen Konzepten (z. B.: *two eyes*) als bei einem Pluralzahlwort mit konzeptuellem Dual (z. B.: *eight eyes*). Bei den Muttersprachlern ist das Ergebnis eindeutig. Der konzeptuelle Dual in Kombination mit dem Zahlwort “zwei” zeigt einen deutlichen Verarbeitungsvorteil, der auch den *Subitising-Effekt* aufhebt. Die konzeptuelle Information beeinflusst also sehr klar den Verarbeitungsprozess, auch wenn die Probanden sich auf eine grammatikalische Kongruenz konzentrieren. Die Lerner der Klasse 7 und der Klasse 11 zeigen die gleichen Resultate wie die Muttersprachler. Bei Klasse 9 sind die Unterschiede fast signifikant, aber bei Klasse 10 und 12 gibt es keine Unterschiede. Bei den Lernern sind die Ergebnisse also nicht so eindeutig. Die Anzahl der Lehrjahre für das Fach Englisch scheint bei den Ergebnissen wieder einmal keine Rolle zu spielen. Bedeutungsvoll ist, dass sich Klasse 7 und 11 wieder von den anderen Gruppen hervorheben, wie bereits beim Auftreten des Numerus-Effekts. Da angenommen wird, dass konzeptuelle Prozesse sprachunabhängig sind, wurde ein Effekt auch bei den Lernern erwartet. Sprachlerner könnten jedoch eine Sprachfertigkeitsschwelle besitzen, die sie überschreiten müssen bevor diese konzeptuellen oder auch probabilistische Effekte eine Rolle spielen. Diese Effekte sind zwar in der Muttersprache vorhanden, eine Hypothese wäre aber, dass die Verarbeitungslast beim Gebrauch der Fremdsprache zu groß ist, als dass diese Effekte noch eine Rolle spielen könnten. Die Tatsache, dass bei einigen Lernern diese Effekte durchaus auftreten, stützt die Annahme, dass sie eine psychologische Relevanz besitzen und messbar sind, wenn sie nicht durch andere Verarbeitungslast blockiert werden.

Schlussbemerkungen

Die Ergebnisse der beiden Experimente belegen deutlich, dass extra-linguistische Faktoren bei der Sprachverarbeitung eines so unauffälligen Flexionsmorphems eine große Rolle spielen. Probabilistische Faktoren wie Frequenz und Dominanz haben eine große Relevanz bei der Sprachverarbeitung, aber auch konzeptuelle Kongruenz, wie aus dem zweiten Experiment klar wurde. Die Sprachverarbeitung eines Flektionsmorphems, das vordergründig so eindeutig und "einfach" wirkt wie das - *s* des Pluralmorphems im Englischen, ist in Wirklichkeit ein komplexes Zusammenspiel einer Vielzahl von Faktoren. Bei der Sprachverarbeitung ist es also nicht möglich, strikt zwischen grammatikalischer Regelanwendung und Zugriff auf das lexikalische Gedächtnis zu trennen. In anderen Worten, man kann nicht einfach davon ausgehen, dass alle Singularformen im Lexikon gespeichert sind und deren Pluralformen durch grammatikalische Regelanwendung gebildet werden. Es gibt eher einen Grad der Pluralität, der den Verarbeitungsablauf bestimmt. Dieser Grad der Pluralität, für den es ein großes Kontinuum gibt, wird durch eine Vielzahl von Faktoren wie Frequenz, Dominanz, Interaktion von Frequenz und Dominanz sowie konzeptuelle Faktoren bestimmt. Die Ergebnisse der durchgeführten Experimente zeigen, dass nicht von einer strikten Trennung zwischen Grammatik und Lexikon bei der Sprachverarbeitung ausgegangen werden kann.

Eine weitere wichtige Beobachtung lässt sich beim Vergleich zwischen Muttersprachlern und Sprachlernern feststellen. Allein die Trennung dieser beiden Probandentypen zeigt, dass man in Bezug auf die Sprachverarbeitung große Unterschiede erwartet. Interessanterweise sind diese großen Unterschiede nicht so sehr zwischen den Muttersprachlern und den Lernern festzustellen, als vielmehr zwischen den einzelnen Gruppen innerhalb des gleichen Probandentypen, also Lerner oder Muttersprachler. Bei den Sprachlernern ist keine graduelle Verbesserung oder Beschleunigung der Reaktionszeiten von Gruppe zu Gruppe zu erkennen, die eine Entwicklung hin zur muttersprachlichen Verarbeitung aufzeigt. Es scheint, dass der Sprachinput im formalen Sprachunterricht nicht ausreicht, um hier große Entwicklungsstufen bei der Verarbeitung zu verzeichnen. Dies ist eine ziemlich ernüchternde Erkenntnis für Sprachlehrer, die ja sicherlich davon ausgehen, dass längerfristig die explizite Sprachkenntnis auch implizit wird und bei der Sprachverarbeitung sichtbar werden müsste.

Auch die beiden Gruppen der Muttersprachler zeigen Unterschiede in ihrer Performanz. Es gibt große Verarbeitungsunterschiede zwischen den beiden Muttersprachlergruppen; in einigen Fällen sind diese Unterschiede sogar größer als zwischen Muttersprachlern und Lernern. Die beiden Experimente verdeutlichen daher auch die große Bedeutung des Aufgabentyps und der täglichen Routine für die Art der Sprachverarbeitung. Die Soldaten haben in ihrer täglichen Routine vermutlich ganz andere Voraussetzungen als die Schüler, die es gewohnt sein mögen, auf schriftliche Stimuli auch am Bildschirm zu reagieren. In Bezug auf die typischen Verhaltensmuster der Sprachverarbeitung bei Muttersprachlern oder Lernern, scheint es in meiner Studie keinen Hinweis dafür zu geben, dass es einen “typischen” Muttersprachler und einen “typischen” Lerner gibt. Daher denke ich, dass auch bei der Verarbeitung ein großes Spektrum an Möglichkeiten besteht, das unabhängig von Muttersprachlern oder Lernern ganz individuell durch die linguistischen und extra-linguistischen Faktoren, die bereits genannt wurden, bestimmt ist. Die Konditionen, unter denen die Probanden arbeiten, wie z.B. tägliche

Routine, Vertrautheit mit dem Aufgabentyp, usw. scheinen auch hier ausschlaggebend zu sein.

Diese Erkenntnis mag wiederum für die Sprachlehrer hoffnungsvoll klingen, da Verarbeitungszeiten kein so verlässliches Maß für Sprachentwicklung sein können. Eindimensionale Verarbeitungszeiten berücksichtigen nicht die unterschiedlichen Dimensionen, die Verarbeitung beeinflussen. Eine mögliche Konsequenz für den Sprachunterricht dürfte somit auch klar sein: er muss die verschiedenen Dimensionen berücksichtigen. Es hat sich gezeigt, dass Sprachlernen nicht nur aus einem reinen lexikalischen abspeichern und grammatikalischer Regelanwendung besteht. Vielmehr müssen bei diesem Prozess eine Vielzahl von lebensechten Aufgabentypen und die Bedeutung von probabilistischen Faktoren, die häufiges Wiederholen unerlässlich machen, mit einbezogen werden.