# Bachelor's Thesis
# Crosslinguistic Morphological Priming

Klara Anna Weiand
University of Osnabrück
Student number 909590
kweiand@uos.de

November 21, 2005

First Supervisor: Dr. Ton Dijkstra, Radboud University Nijmegen
Second Supervisor: Dr. Graham Katz, University of Osnabrück

**Abstract**

This study uses a lexical decision experiment with short lag priming and Dutch and German items to determine whether there is a distinctively morphological component in cognate priming.
The hypothesis that there is a morphological component in Dutch-German cognate priming could be verified.
It was also found that orthographical and semantic priming effects were similar in magnitude and nature to the effects found in studies employing only one language and that the priming effect for complex categories are not purely additive.
However the latter findings could not all be established as statistically significant and thus requires further investigation, since the lack of statistical significance argueably could be caused by the fact that the knowledge of German among the participants was relatively bad.

# Contents

# 1 Introduction

The Dutch-German cognates "gelegenheid - Gelegenheit", "lezer - Leser" and "eigenschap - Eigenschaft" are similar on multiple levels.
Obviously they are spelled similarly and if one knows both Dutch and German it is also apparent that they refer to the same things.
But that is not everything that the word pairs share. The words also have a similar morphology, that is, the words in a cognate pair resemble each other in their inner structure, specifically their suffixes.
The goal of this study is to test the theory that there is a distinctively morphological component in Dutch-German cognate priming. Additionally the distinct priming effects of similarity in orthography, semantics and morphology are investigated, allowing both for a comparison with results from similar studies engaging only one language as well as an answer to the question whether composite priming effects are purely additive.
This approach combines two current topics in psycholinguistics, bilingualism and morphological priming. Recent studies have found that the lexicons and systems for language processing in bi- or multilingual speakers are more integrated than was long assumed.
From these findings and the fact that bilingualism is not a rare exception arises a need to approach the study of language processing and modeling of processes not strictly separated by languages but with taking bilingualism into account.
The main difficulty in this plan is to find an experimental setup that allows to reliably distinguish and control for similarities. This controlling of levels of similarities happens through the material selection where words for all categories in the experiment must be found.
This study uses a lexical decision experiment with short lag priming. Only native Dutch speakers with knowledge of German were tested.
The hypothesis that there is a morphological component in Dutch-German cognate priming could be verified.
It was also found that orthographical and semantic priming effects were similar in magnitude and nature to the effects found in studies employing only one language and that the priming effect for complex categories are not purely additive.
However the latter findings could not all be established as statistically significant which is assumed to be caused by the fact that the knowledge of German among the participants was relatively bad and thus requires further investigation.
The theoretical foundation and previous results from related studies are presented in section 2. The third section is concerned with the design, planning

and execution of this study.
The results of the experiment are presented in the fourth section. The fifth section finally discusses the outcome of the experiment.

# 2 Theory

## 2.1 Introduction

This section introduces basic concepts in morphology and gives a short overview over the role of morphology in psycholinguistics and visual word recognition.

**Morphology** Morphology studies the structure of words and their constituents, morphemes.
Morphemes are the smallest meaningful units of language.
Linguists distinguish between free morphemes and bound morphemes.
Free morphemes are morphemes that can stand on their own and thus constitute whole words. Bound morphemes on the other hand must be combined with other morphemes in order to occur in a sentence.
Morphology is productive in that there are rules which allow for the formation of new words from morphemes or other words through combination and alteration.
In the well-known "wug" test[2] children were presented with a picture of an animal and were told that it was called a "wug". When shown a picture with two of the animals and asked to complete the sentence "There are two _____", most children as young as six years have no problem finding the "correct" answer, wugs[1].
This result illustrates both that at least not all words are stored in the mental lexicon as a whole (with "wug" being a non-word, it was clear that the children could not have learned its plural before), without regard to morphology, and that morphology has productive rules.
A less obvious conclusion that can be drawn from the example of plural formation is that the meaning of a word is combined from the meaning of its morphological constituents. While "wug" refers to a single entity, "wugs" refers to two of them.
Thus, just like adding the plural morpheme alters the syntactic properties of a word, its meaning is changed, too. This becomes even more obvious in derivational morphology where the meaning of a root is usually changed more than it is by inflection:
Inflection is the modification of a word such that it displays grammatical information (e.g. adjusting number and tense) while derivation is concerned with the formation of new words from already existing words[4]. While words never change their word class in the process of inflection, they often do in

[1] the matter of allomorphs is set aside here

derivation.
Another difference is that while the change in meaning of an inflected form is predictable, the result of a derivational process does not always succumb to the same semantic scheme.
For example German "Xität" often refers to a state of being X (the same is true for Dutch "-iteit") but while this works for "Komplexität" and "Musikalität", it does not for "Lokalität" which is not the state of being local and "Extremität" which only in one possible sense of the word is the state of being extreme. Recent theories argue that the distinction between inflectional and derivational morphology is not as clear-cut as it has long been considered. Another issue is whether in language production and understanding there is one system both for derivation and inflection or whether they are two distinct modules.
Since this study is concerned solely with derivational morphology, these issues will not be considered further in the following.

**Morphological Priming** In the past 25 years many aspects of morphological processing have been investigated. One of the most fundamental questions is whether there is a uniquely morphological component to visual word recognition or whether, as some theories, many of them employing connectionist methods, state, there is no morphological level of processing that can be discriminated from semantic and orthographic or phonetic components.
Facilitory morphological effects on reaction time and error rates have been demonstrated under a variety of conditions[31]: In lag priming, effects in the absence of semantic effects have been shown in Hebrew[1] and in the absence of orthographic effects in German[12] and English[37].
They also have been shown in masked priming in the absence of semantic effects in Hebrew and in the absence of orthographic effects in French[21], Dutch[12], American Sign Language[25] and Serbo-Croatian[31].
The existence of a separate morphological level of processing has been shown in different languages (Finnish[26], French[32], Spanish[11], Hebrew[19]) and not only in visual word recognition but independent of modalities[32], that is, independent of whether prime and target were both presented visually or auditorily or whether one was presented auditorily and one visually.
The goal of this study is to investigate whether morphological effects can be found in crosslinguistic priming using two languages with a similar derivational morphology, Dutch and German.

## 2.2 Methodology

In this section, the different aspects and factors of experiment design for studying morphological, orthographic and semantic factors in visual word recognition are discussed. Along the way, possible problems for experimental designs in such a context are discussed.

**Priming** Priming is often used in order to study the component effects in visual word recognition. The idea behind priming is that once certain connections or associations have been used, they have a different activation threshold for a certain amount of time; an acuteness to certain kinds of stimuli or features is thus established.
Priming is assumed to be an involuntary (and thus automatic) and often unconscious phenomenon since it occurs even in tasks where no information about the previous trial is required or useful. This sets it apart from deliberate retrieval as is needed in episodic memory tasks which, for example, require conscious retrieval of elements on a list presented to the subject.
In *identity priming* a whole word or some of its aspects are repeated over the prime (the first presentation of which establishes the sensitivity) and the target, the second presentation. The effect of the priming is then measured compared to a control condition which determines the baseline relative to which the effect of priming is established.[16]
The recording of priming effects is usually operationalized by measuring reaction times to certain tasks involving the prime and target. Another possible factor are error rates, measured in percent.[33]
Priming effects can be inhibitory or facilitatory, that is, the reaction time can increase or decrease compared to the baseline as a consequence of priming. If there is no priming effect, the condition tested does not differ significantly in terms of reaction time from the control condition.
An example: Priming with an identical word can have a facilitatory effect lasting up to 48 hours or more, that is, if a word is presented to a participant in a task and again up to two days later, the reaction time on the second presentation will be significantly shorter than that for a word that has not been presented before[41].

**Priming Paradigms** For the experiment design not only the different possible tasks are important but also the priming paradigm used. The goal of priming is to show an effect that is free from influence through external or unwanted factors and only reproduces the effect that is studied. This must be ensured not only through careful selection of the material and material

conditions but also through the experimental setup and technique.
The main focus is on eliminating strategic and episodic effects[16].
Episodic effects are effects that are generated by a memory of the specific perceptual aspects of an episode, that is, a presentation of an item. So, rather than reflecting only the activation the lexical representation of the word presented, the reaction time may also be influenced by the memories of the presentation of the related or identical item.
An extreme example of a strategic effect on the other hand would be a lexical decision task where all letter strings are words. The participants would of course notice this soon and either employ the strategy of giving a positive response independent of the letter string presented or at least develop a bias towards a positive response.
But of course strategies can not only be used if the proportion of non-words and words is not balanced or if their arrangement is so regular that a pattern can be deduced.
In general it is important to prevent the participant from becoming aware of any regular relations or patterns in and between the strings he or she has to make a judgement about. This goes both for the relations between words that are to be investigated and other regularities that are not the focus of the study, because any strategy –even a faulty one– can confound the results of the experiment.
So the goal of the different priming paradigms is to minimize strategic and episodic effects in order to guarantee a recording of lexical effects that is as pure as possible.
One priming paradigm often used is masked priming where the prime is presented –usually for an amount of time so short it cannot be consciously detected– sandwiched between a forward and a backward mask (######).
Strategic and episodic components are eliminated insofar as, if the participant does not consciously perceive the prime, he or she cannot develop a strategy or form an episodic memory of the event of seeing it.
Another priming paradigm is lag priming where prime and target are separated by a number of items, the "lag", sometimes as many as fifty in long lag priming and as few as one or two in short lag priming. One big difference compared to masked priming is that in lag priming the prime is presented as a regular item and e.g. in a lexical decision experiment the participant makes a lexicality judgement about the prime as well as about the target which can have an influence on the results.[16]
The two priming paradigms just discussed display different approaches to minimizing strategic effects:
In masked priming the prime is presented for a very short time which is assumed to limit its processing and in this way prevent conscious appreciation.

Lag priming on the other hand relies on varying lags. It has been shown that participants can detect relationships between primes and targets even if they are separated by several intervening items, so lag length alone cannot be responsible for preventing interference from unwanted effects.[16]
The key to this in lag priming is to keep the variance in the items high. Not more than half of the items presented should be words and the number of items between primes and targets should differ and be well (i.e. pseudo-randomly) distributed. Additionally, prime-target pairs should not all be related in the same, easy to detect and obvious manner and the distribution of prime-target pairs from different condition should also be pseudo-random.
In the study of morphological priming effects masked priming and lag priming show similar results which indicates that both techniques can be equally good when it comes to preventing strategic and episodic effects.

**Tasks** One of the tasks used most often in visual word recognition is that of lexical decision. In a lexical decision task the participant is presented with a letter string and must indicate whether that letter string constitutes a word of the language in question as quickly and accurately as possible[16]. The idea behind the task is that the reaction time and error rate of reacting to a word stimulus reflects how quickly and efficiently the representation of the word in question can be activated.
The stimulus can be presented in different modalities, auditorily or visually or mixed, and usually the lexicality judgement is indicated by pressing one button for "yes" and another one for "no". There are also experiments, where a different reaction is expected from the experiment, e.g the letter string has to be read out loud if it is a word.
Another possibility are conceptual tasks where conceptual judgments about the referent of a word have to be made, for example the participant has to indicate whether the words presented to him represent living objects or not. Since lexical decision tasks are considered to rely more on lexical processes, conceptual tasks are often employed in experiments where the stress is put on semantics and where conceptual processing is to be encouraged.
Yet another task is segment-shifting where participants have to move an underlined part of one word to the back of another word and pronounce the result.

**Distinguishing Levels** The main problem, however, is that, due to the compositionality of morphology, similarities on one level often entail similarity on another level. This makes the different factors, form, semantics and

morphology, hard to discriminate.
In the case of morphology this problem exists both for root morphemes and affixes, both through which morphological relatedness can be defined. Although the surface form or individual representation of morphemes may differ, they usually share orthography and, depending on language, phonology as well as semantics. This covariation lends itself to orthographic/semantic accounts of morphological effects[31].
This is why finding a way to discriminate the different aspects of a word is a major challenge for showing a distinctively morphological component in visual word recognition.
As relatedness on the different levels often occurs with relatedness on others, it is crucial to not only control relatedness in the relevant conditions but also ensure the absence of relatedness in the other factors of the condition (the configuration of values of relatedness on the level of form, semantics and morphology) to prevent unintended influence from other aspects.
For example, "CAR" and "CARD" are only orthographically related but "CAR" and "CARS" are orthographically and morphologically related.
However, the second pair of words is also very closely semantically related which is why a significant difference in reaction times for the two prime-target pairs in an experiment does not allow for a conclusion about the existence of morphological priming.
It also has to be noted that it may not be enough to show separately that morphological priming effects cannot be explained in terms of orthography (comparing priming effects from morphologically and orthographically related word pairs to word pairs only related in terms of orthography, e.g. "DIE" - "DIES" and "DIE" - "DIET") and that they cannot be explained in terms of semantics[23].
This is because this leaves open the possibility that orthographical effects and semantic effects interact in a way such that what is perceived to be a morphological priming effect cannot be explained with effects from orthography or semantics alone but can be explained as an interaction or summation of orthographic and semantic effects that, on their own, may not be significant.
That means that, since levels of similarity are so hard to vary independently of each other, word pairs with a similar orthography and similar morphology may have some kind of semantic relation, either through a common suffix or root morpheme, which might cause a non-significant semantic effect that in interaction with the orthographic effect might account for what is seen as a morphological effect.
Often studies only compare effects from orthographically or semantically related word pairs with word pairs that are additionally morphologically re-

lated.
So it is important to find conditions where morphological effects can be discriminated both from orthographic and semantic effects as well as from the combination of the two and where the combination of similarity values for the three different factors can be varied and fully controlled. The configuration of values that offer this is and which both very useful and problematic because examples are hard to find is word pairs that have orthographic overlap and similar semantics but that are not related morphologically.
This condition is important because it is needed to explore the possibility mentioned above, the idea that morphological priming is actually just a combination of orthographic and semantic effects. This possibility can be excluded if there is a significant difference in reaction times between word pairs that are orthographically and semantically similar and word pairs that additionally share morphology.
An approach to this is to use onomatopoetic[2] words, portmanteaus[3] or phonaesthemes[4] because those kind of words display a non-arbitrary mapping between form and meaning without a morphological component[36].

**Semantic Transparency** If morphologically similar words that contain the same root morpheme are considered, semantic transparency is another important variable. Semantic transparency describes the degree to which morphologically related words are similar in their meaning. The meaning of semantically opaque words can thus not be predicted from the meaning of their constituents.[8]
In a cross-modal priming experiment focusing on derivational morphology, Marslen-Wilson, Tyler, Waksler and Older [30] found a facilitatory morphological effect only when prime and target, one of which was the unaffixed root morpheme, were related in a semantically transparent way but results from several other studies differed in that the semantic transparency had no influence on the morphological priming effect[29].

[2]words that sound like the action they describe, e.g. "click", "fizz"
[3]words that are formed by combining the form as well as meaning of two other words, e.g. "smog" ("smoke" + "fog") and "motel" ("motor" + "hotel")
[4]semantically related words that share part of their phonetic structure

## 2.3 Previous Work

**Findings** Most research on morphological priming is focused on showing the existence of a morphological effect that is independent of both orthographic and semantic effects as well as their combination.
It has been shown in various studies that both in masked priming and lag priming orthographic priming is inhibitory while morphological priming is facilitatory[31]. Semantic priming on the other hand is facilitatory but extremely short-lived, again, which has been found both in masked priming and lag priming[1], where it was significant only after a SOA (stimulus onset asynchrony) of 1 second but not at 4 seconds and after no intervening items but not after 15.
One of the few time course studies of morphological priming by Rastle, Marslen-Wilson and Tyler[36] investigated orthographic, semantic and morphological priming effects at SOA times of 43, 72 and 230 ms. Words were said to be morphologically related if they contained the same root morpheme and semantic relatedness of word pairs was rated by participants on a scale of 1 to 10 in a separate experiment.
Rastle et al. found that orthography on its own had a small, nonsignificant inhibitory priming effect while semantically similar word pairs showed a similar pattern as the orthographically and morphologically related word pairs, namely a facilitatory effect that is only significant at the SOA of 230 ms.
The "cognate" condition on the other hand displayed an early (43 ms) significant facilitatory effect that declined with growing SOA.
Their results indicate that there is a uniquely morphological component that cannot be reduced to an addition of orthographic and semantic effects. It also suggests that morphological processing takes place early in the word recognition process.
Feldman and Protsko[15] on the other hand found a morphological effect that increased with SOA, however they did not control for all aspects in each condition of their experiment (e.g. VOWED-VOW was considered morphologically similar but the word pair displays similarities on the orthographical and semantic level, too) which makes their results less clear and harder to interpret.
Confirming the observation of little impact of orthography, Fowler, Napps and Feldman[17] found that the facilitatory effect of morphology is only slightly diminished for semantically and morphologically related word pairs compared to word pairs in the cognate condition in lag priming.
Another approach to investigating morphological effects that does not include lexical priming considers root frequencies:
Since words are recognized faster if their frequency is higher, it is assumed

that word frequency is incorporated in the mental lexicon.
Several studies have shown effects of cumulative root frequency[21], that is, not the frequency of the whole word but that of its root morpheme has an influence on the speed of word recognition. This suggests a morphological decomposition in the process of word recognition.
Seidenberg and Gonnerman on the other hand state that this is too strong a conclusion[22] since cumulative root frequency effects appear in distributed connectionist network models that do not have an explicit level of morphological processing such as the Seidenberg and McClelland model.

**Models** Basically, there are three approaches to modeling lexical access in visual word recognition that incorporate a morphological level[8][38]: single route, dual route and supralexical models.
In single route models a form-based morphological parsing mechanism strips word off their affixes thus isolating the root morpheme. Since the parser works on the orthographic form, it does not work for exceptions. For example, removing a word-final "er" works well for "baker" and "finder" (or "Bäcker", "Finder" and "bakker", "vinder" in German and Dutch) in that removing it yields the root morpheme, but no roots exist for "finger" and "grocer".
A big disadvantage is that the parser will be wrong in every case where a character string that is identical to an existing affix appears in the correct position without actually being a morpheme.
Tafts and Forster[38] found that rejecting non-words is more difficult if they consist of an actual stem with an affix that exists but cannot be attached to it than when they consist of the same affix and a nonexisting stem.
On this basis it was proposed that stem morphemes are the access codes employed in lexical lookup. However, it may be argued that non-words may not reflect the normal process of word recognition and that decomposition might be a process only resorted to when a lookup of the representation of the whole word does not find a matching entry.
Dual-route models encompass this interpretation. They do not only possess the decompositionary route of the single router models, but also a routine that tries to find the whole word in the lexicon. For example in race models the two routes compete and dmssbthe slower or unsuccessful one consequently plays no role in the recognition of the word.
An alternative is the supralexical model in which morphemes are only activated after the representation for the whole word form has become active. The morpheme representation then sends back activation to the words that

are compatible with it. In this model morphemes thus are abstract, amodal underlying representations that are learned through regularities[8].

**Alternative Accounts** As already mentioned there are several theories that try to explain morphological priming effects through orthographic or orthographic and semantic effects, the main supporters of these theories being Seidenberg and Gonnerman[23][22].
Seidenberg proposes that patterns of high and low bigram frequency could account for so-called morphological effects and that sublexical structure in word recognition thus is only an emergent phenomenon since the bigram probabilities at morpheme and syllable transitions are lower than the sequence probabilities within those units. This however is not consistent with the finding that the priming effects from only orthographically similar word pairs and orthographically and morphologically similar word pairs are different.
In a later paper[22], Seidenberg and Gonnerman question what they call "discrete morphology", the view that words consist of morphemes whose combination is governed by rules and instead propose "an approach to thinking about complex words".
They remark that not only is there a vast number of words that do not conform to the ideal of classic morphology -discrete units with a combinatorial semantics and clear rules- but also that the criteria for classifying words as complex or not are unclear and subjective which makes empirical results vague and hard to interpret unambiguously. Instead of a rule-governed system with exceptions they propose a connectionist network where graded effects from relations between orthography, phonology and semantics replace the notion of discrete morphemes and combinatorial rules.
While Seidenberg's and Gonnerman's criticism of the lack of a proper ontology for classification of morphological relatedness[5] and ensuing subjectiveness and vagueness is legitimate, their suggested model also is not compatible with the empirical finding that words that share orthography and semantics show a significantly different effect than words that are orthographically as well as semantically and morphologically related. It seems that no reductive theory can hold given that this finding is valid.
Another argument against reductionist accounts comes from a long lag priming lexical decision experiment in Serbo-Croatian[14] where target and prime were presented in different alphabets (Roman and Cyrillic) and compared to the results of another experiment where the only difference was that primes and targets were written in the same alphabets. No significant effect from

[5] if such a thing is possible

the alternation of alphabets was found which was interpreted to show that orthographical similarity is no precondition for morphological effect.
This argument is not as strong as the first one since the findings are compatible with a theory where morphological effects are explained through semantics and phonology instead of semantics and orthography since, while orthographic form differs between the conditions, phonological form does not. Also, one could postulate common representations for equivalent letters in different alphabets and consequently and on this basis hold on to theories that reduce morphological effects to semantics and orthography.

## 2.4 This Study

The goal of this study is to extend the research on morphological priming to a crosslinguistic perspective. Short lag priming is used in a lexical decision experiment in order to investigate orthographic, semantic and morphological effects.
The impulse for this comes from the insight that bilinguals do not possess two completely independent word recognition systems[9]. This was assumed until about twenty years ago but more recent findings have shown that bilinguals do not possess two completely distinct lexicons that can be switched on and off if needed.
For example Dijkstra, Timmermans and Schriefers[10] executed a lexical decision experiment with Dutch-English bilinguals who had do press a button if the word presented on the screen was English and wait if it was Dutch. They found that Dutch-English bilinguals took significantly longer to answer when the word presented was a Dutch-English homograph, that is, words that are orthographically identical but have different meanings.
All this indicates that the participants activated both their English and Dutch lexicon, although only the English one was needed to successfully master the task.
But influence on visual word recognition can not only the go from the native language to the second language, the influence can be in the other direction, as well.
Bijeljac-Babic, Biardeau and Grainger[3] did an experiment involving masked orthographic priming with French-English bilinguals. They demonstrated that in a lexical decision task the inhibitory orthographic effect commonly found in orthographic priming was not only present when both prime and target were French words. Rather, an inhibitory orthographic priming effect was also found when the prime was an orthographically related English word. The strength of this effect depended on the participants' proficiency in English.
Similarly, Dijkstra et al.[10] found a homograph effect when they repeated their experiment described above switching the role of Dutch and English in the task. That is, Dutch-English bilinguals had to react only if the word presented was Dutch.
Influences of both languages on visual word recognition can be found in monolingual experiments which distinguishes a bilingual from a monolingual even in monolingual tasks involving their common native tongue.
Because of this and since bilingualism (or even multilingualism) is rather the norm than an exception, theories and models of visual word recognition should not be restricted to one language but should much rather be able to

account for second language processing, as well.
These considerations give rise to the question how the influence of other languages found in visual word recognition can be accounted for in terms of theories and models of bilingual lexical organization and access.
When it comes to the general framework of bilingual lexical representation, there are two possibilities; Either bilinguals possess one integrated lexicon for both languages or two distinct and independent lexicons.
On the other hand lexical activation may either be language-selective or not, that is, language selection could be accomplished during the process of lexical activation or before it.
Combining those two factors, four different possible models emerge: Language-specific lexicons with language-selective or non-selective access mechanisms or integrated lexicons with language-selective or non-selective lexical access[39].
Van Heuven, Dijkstra and Grainger[40] examined the influence of orthographic neighborhood size in an experiment with Dutch-English bilinguals. The orthographic neighbors of a word are words that differ from it only by one character. It was shown that the number of orthographic neighbors in one language had a significant effect on the reaction time to a target word in the other language. This indicates that the bilingual lexicon integrates both languages and has an access mechanism that is not language-selective. All other three possible models cannot account for the results since there is either no interaction between the lexica or because not both languages are active at the same time.
So there is evidence for a non language-selective lexical activation process that operates on an integrated lexicon for both or all languages. These findings both provide the basis for this work and justify the need for investigating crosslinguistic morphologic effects.
If lexical activation and lexical organization were found to operate strictly in isolation for each of the languages of a bi- or multilingual, there would be no reason to assume that the case would be different for morphology, that is, that morphological activation in one language would have any influence on the performance of morphology in another language.
But since the opposite is the case, the question arises whether uniquely morphologic effects can be found across languages and in what relation they stand with orthographic and semantic effects. The studies mentioned above thus constitute the foundation for this study.
The study by Bijeljac-Babic et al.[3] mentioned before found an orthographic effect in a crosslinguistic task that was very similar in nature to that found in experiments employing only one language.
The case for semantic priming effects is less clear due to conflicting results both in monolingual and crosslinguistic priming.

While some studies found facilitatory semantic effects for synonym priming in lexical decision tasks, other studies were unable to detect a significant effect. The same is true for crosslinguistic lexical decision experiments using priming with translation equivalents[24][28][42].
However, most of these studies agree that semantic effects are not very strong (or even not present) in lexical decision tasks and more apparent in tasks that demand more conceptual processing[42].
The goal of this study now is to show that there are not only orthographic and semantic effects in crosslinguistic lexical decision tasks, but also morphological effects.

**Cognate Suffixes** Most studies so far consider words that contain the same root morpheme (e.g. "DEPART - DEPARTURE") to be morphologically similar. Since this study is different from all studies mentioned in that it deals not with morphological priming effects within a language but across two languages, Dutch and German, it depends on different assumptions than the monolingual studies and of course the methodology must be adapted accordingly.
If this operationalization was was used in this study, the focus would not only be on whether there are common morphological processes for Dutch and German, but also bigger issues about the organization of the bilingual mental lexicon would arise. Translation equivalents of German and Dutch root morphemes differ not only in their orthographical and phonological representation, but also partly and subtly in their semantics and syntactic frame. Because of this, the focus here is put on the morphological processes by defining morphological relatedness as two words bearing cognate suffixes. Since there is no common taxonomy for classifying suffixes as cognates, in the following an attempt is made to formulate criteria for determining whether two suffixes are cognate suffixes.
Cognates play a role in many areas of linguistics and are often referred to in historical linguistics as well as in sentence alignment and cross-linguistic studies. There is a huge number of different, sometimes conflicting definitions:
All definitions agree that cognates are words which share similarity on some level but that is as far as the unity goes. Factors are orthographic and phonological similarity, a common etymology and similar semantics.
For example Duyck, Diependale, Drieghe and Brysbaert[13] and Ziegler, Perry, Ma-Wyatt, Ladner and Schule-Koerne[43] consider translation equivalents and a very similar but not identical orthography to be "near cognates" which suggests that they see cognates as translation equivalents with identi-

cal orthography.
Piao, Clough and Arundel[35] on the other hand adopt Simard's definition according to which cognates are words from different languages that have an overlap in orthography and semantics.
Crist[6] explicitly contradicts both these approaches saying that cognate status does not have to do with similarities in form and meaning but solely with whether two words in different languages descend in form from the same word and whether the two languages are related.
This study strikes a balance between the different theories presented assuming that two suffixes are cognates if they tend to appear as constituents of cognate words, have a similar or identical orthographic form and semantics and, if it could be determined, a common ancestor.
An additional constraint that takes into consideration the role of morphological structure is that cognate suffixes should attach to mainly words of the same classes and yielding words of the same word classes.
As an example, Dutch "-BAAR" and "-BAR" only differ by one letter in form and both attach to transitive verbs resulting in an adjective. They probably both are related to Old High German "-BARI" and "V-bar" and "V-baar" often mean "can be V-ed" as in Dutch "LESBAAR" and German "LESBAR" (readable) and "ZICHTBAAR" and "SICHTBAR" (visible).

| **Yields** | **Dutch** | **German** |
|---|---|---|
| ADJ | baar | bar |
| N | dom | tum |
| ADJ | en | e(r)n |
| N | er | er |
| N | heid | h\|keit |
| ADJ | ig | ig |
| N | ij | ei |
| N | ing | ung |
| N | iteit | ität |
| ADJ | lijk | lich |
| ADJ | loos | los |
| N | schap | schaft |
| ADJ | zaam | sam |

Table 1: cognate suffixes used

# 3 Methods

## 3.1 Design

The experiment used lag priming with either no or one intervening item between prime and target in a lexical decision task. This means that a reaction was required both for the prime and the target.
The response was indicated by pressing buttons on a button box and the reaction time was measured in milliseconds.
To investigate the influence of orthographic, semantic and morphological relatedness, seven different categories of word-pairs were considered.
Word-pairs were classified according to their orthographic, semantic and morphological relatedness. In the following the three aspects are abbreviated as "o", "s" and "m" and the presence or absence of relatedness is indicated with "+" and "-". A set of "o", "s" and "m" together with an assignment of a binary value for each factor is referred to as a condition or category, a total of which there are eight.
So for example "+o-s-m" denotes word-pairs that only have similar orthographical representations but that are not related semantically or orthographically. Controlling the absence of relatedness as well as its presence ensured that there could be no unwanted influence from uncontrolled factors.
The categories used in the design and execution of the experiment were "+o-s-m", "+o-s+m", "-o+s-m", "-o+s+m", "-o-s+m", "+o+s-m" and "+o+s+m". This set-up allows to investigate single effects from orthography, semantics and morphology as well as their interaction.
Relying on the *ceteris paribus* assumption we conclude that, if two conditions are identical in two factors and only differ by one factor, a difference in outcomes (i.e. in this case, a significant difference in priming effects between two categories which can be observed through a difference in the average reaction times to the prime and the target) is to be attributed to the one differing factor.
This allows for a consideration of components of effects in isolation and for the investigation of interaction between different effects.
The categories were chosen such that they permit a look at all three effects in the absence of the other two effects. Additionally, all three possible combinations of two effects were used as conditions in order to investigate the interaction between effects. The final condition is the cognate condition where word pairs are related orthographically, semantically and morphologically. This setup makes it possible to consider the single effects as well as the combination of effects and the interaction between them.

## 3.2 Material Selection

The classification of the stimulus material was done using a script written in the scripting language Ruby[6] that applied the constraints described in the following to all combinations of Dutch-German word pairs. These were retrieved from the CELEX[7] morphological lexicon files for the corresponding language which contain several different illustrations of a word's morphology as well as information about frequency and word class.
All words annotated with other pos-tags than adjective and noun were filtered out and excluded from further consideration since bimorphemic suffixed words were to be the focus of the study.
The set was restricted to those to word-classes since simple proper nouns generally are not morphologically decomposable in the sense that they have semantic transparency. On the other hand verbs were left out because they are often prefixed and thus do not conform to the structure we are looking for.
German frequencies were extrapolated since the German corpus from which the frequencies were derived was far smaller than the Dutch one and words with extremely low frequencies (less than three per million) were removed from the data set.
In the next step, all possible Dutch-German noun-noun and adjective-adjective pairs were checked for their difference in word length and frequency and excluded if they were too big. That was the case when their lengths differed by more than one character or if the absolute value of the difference in frequencies was bigger than the frequency of the Dutch word in the pair. This was done in order to avoid unwanted effects stemming from those two factors.
After that, the script checked whether the words left actually had the right structure and consisted of two morphemes, the root and a suffix.
The word pairs that were now left were classified according to three binary features: orthography, semantics and morphology.

**Orthography** The Levenshtein or Edit distance was used as a means to quantify and operationalize orthographic overlap. The distance of two strings is the number of operations minimally needed in order to transform one into the other. Feasible operations are insertion or deletion of characters and the substitution of one character with another one[27].
As an example, the edit distance of the Dutch SCHEPPER and the German SCHOEPFER is 2 as the shortest solution to get from the first word to the

[6] http://www.ruby-lang.org
[7] http://www.ru.nl/celex/

```
---
2166: ademloos
3809: atemlos
German Frequency:259
Dutch Frequency:262
Distance: 2
German Structure Type: c
Dutch Structure Type: c
---
```

Figure 1: Example output of the ruby script for preselection

```
8656\dankbar\179\C\1\Y\Y\Y\dank+bar\Vx\N\N\N\\
((dank)[V],(bar)[A|V.])[A]\N\N\N\N\I\N
```

Figure 2: Example data from the CELEX morphological lexicon files

second is to substitute the second "P" with an "F" and to then insert an "O" after the "H".
For the computation of the edit distance only the "pseudo root" of each word, which was computed by removing the suffix from the word [8], was considered since otherwise "+o" would entail "+m" due to the orthographic similarity of cognate suffixes (or rather the orthographic dissimilarity of non-cognate suffixes).
The Levenshtein distance determined based on the pseudo root had to be either 1 (if both pseudo roots had the same length) or 2 (if one of the roots was one character longer) for a word pair to be labeled "+o". Words whose orthographical representation was identical were left out since they cannot be unequivocally recognized as a Dutch or a German word which might have an effect on reactions times and thus yield unwanted effects in the experiment. The average distance in the material used was 1.69 for the "+o" condition and 4.73 for the -o condition.

**Semantics** A word-pair was labeled "+s" when one was the translation of the other which was determined by a simple dictionary look-up. All other aspects of semantic similarity were treated manually.

[8]What I call "pseudo root", the practical realization of the root morpheme in a word, sometimes has been called "base morpheme". But since "root morpheme" and "base morpheme" are also often used interchangeably, I prefer to use the term "pseudo root" to avoid confusion

**Morphology** For the assignment of values for morphology cognate suffixes were considered. Just like cognate words, cognate suffixes share a common origin and etymology and usually have identical or similar semantics and phonological or orthographical representations.
Two words were considered morphologically similar if their suffixes were cognates, otherwise they were labeled "-m".
It is important to notice that this notion of morphological similarity does not necessarily entail that two words labeled "+m" possess the same morphological structure although in practice this is often the case since many cognate suffixes are productive only with one word class, the same one in Dutch and in German.

The output of the script were eight files - one for each combination of orthography, semantics and morphology values - containing the corresponding word pairs and their frequencies. The pairs with "-o-s-m" were discarded.
For the selection of the words that would be in the experiment further work was needed. As it was safe to assume that most or all of the participants in the experiment would know English, word pairs where at least one of the words was identical with an English word (e.g. "branding" which means "surge" in Dutch) were eliminated to avoid the coactivation of English vocabulary and potential effects of this, as were word pairs where the Dutch word was also a German word and vice versa (e.g. "monster" which means the same in German and sometimes in Dutch as it does in English but which can also mean "sample" in Dutch).
While the word pairs that had been classified as "+s" needed no further work, it had to be considered that the fact that two words are not direct translations of each other does not mean that they are all semantically unrelated to the same degree and that there are facilitatory and inhibitory effects stemming from the semantic relationship and degree of relatedness between two words[34].
It was not possible to run a separate experiment where participants judge the degree oft semantic relatedness as it is often done or to perform a Latent Semantic Analysis[9], but antonyms, hyperonyms and hyponyms were deleted from the material and close attention was paid to the degree of semantic relatedness within the word pairs during the manual selection of the material for the experiment.
From each category used in the experiment, twenty items were selected and

[9]LSA, a technique for establishing a measurement for semantic similarity through the analysis of co-occurence vectors retrieved from large bodies of text

each item in the target conditions was matched with an item in the control conditions yielding target-control material sets of which each consisted of forty word pairs[10].
For every word pair the difference in length and frequency was calculated and for each set of target and control word pair the differences in lengths and frequencies was determined.
In order to make sure there were no significant differences in the means of the four above-mentioned features in relation to the variation in the data between the material in the different conditions, unpaired t-tests were done for each feature and each combination of two target-control material sets.
This was done to make sure that there were no effects from differences in lengths and frequencies that are not present in all of the categories.
In order to confirm that this is indeed the case, the null hypothesis, that is, the claim that both data sets come from the same population, has to be accepted.
Most textbooks and researchers agree that the level of significance *required* to reject the null hypothesis is p=0.05 which means that we are wrong in 5% of the cases if we assume that the difference in means can be considered coincidental.
However it is not so clear which value p has to take on so it is justified to *accept* the null hypothesis.
While it is very often assumed that if we cannot reject the null hypothesis we have to adopt it, Frick[18] proposes that the null hypothesis should not be accepted for p-values lower than 0.20 and that cases where p is between 0.20 and 0.50 should be considered ambiguous.
The lowest result of all the t-tests run was 0.16 with the majority of p-values being well over 0.5, thus according to the common criterion we can accept the null hypothesis for every feature and all variations of combinations and while this is not completely true for Frick's stricter requirements, there are only very few results that do not unambiguously allow us to adopt the null hypothesis according to him.

**Non-words** 200 non-words were generated from words that fulfilled the criteria for the target and control words but that were not included in the final list. One or two letters were changed to turn those words into non-words. Only the pseudo-root and not the suffix was changed in order to exclude the possibility of using strategies that would lead to unwanted ef-

[10] see "Design"

fects. If all non-words had changed suffixes (e.g. "hoit" instead of "heit"), the participants could classify all presented strings correctly just by looking at the suffix. This of course does not hold for the case that only some of the non-words have altered suffixes, but this also would have allowed the use of strategic approaches, albeit to a lesser extent.
So all non-words had to be seemingly morphologically well-formed with the only difference compared to actual words being that the root word did not exist. Another way to prohibit the use of strategies as described above might have been to create non-words by attaching suffixes to root words yielding nonexistent words (for example because the suffix in question does not combine with the word class of the root word).
However because of the combinatorial semantics of morphology and the high productivity this method would have made the task far harder and maybe even ambiguous without having any advantage compared to using non-word roots.

**The final lists** For the generation of the lists used in the experiment another ruby script was used that randomly picked non-words and word pairs and thus produced lists of 400 words each where the number of items between a word pair could be set as a parameter.
While semantic priming in long-lag lexical decision tasks lasts only a very short time, morphological and orthographic priming has proven to be rather long-lived and can still be shown after a lag of 15 items.
However, since the time course of the priming effects is not considered in this study, a lag of 1 (i.e. the two words of a pair are presented with either one or no other item between them) was chosen so that priming effects would be at their strongest.
The lists generated by the program were segmented into 4 blocks of 100 words and non-words were checked for possible priming effects on the word after them and rearranged if needed. For example the sequence "WECHSELHAFT - RUISTERIG - WISSELVALLIG" was changed to "WECHSELHAFT - BOTTERKEIT - WISSELVALLIG" to avoid morphological priming from "RUISTERIG" to "WISSELVALLIG".
To facilitate later analysis it was made sure that both the time each lag between the words of a word pair was used and the number of times the Dutch and the German word would be the first word of a pair to be displayed were roughly equal.
The resulting lists were then reformatted so they were readable by the pro-

gram used to execute the experiment, a program written at the University for Nijmegen.

## 3.3 Participants

Fifteen participants between the ages of 18 and 53 with the average age being 26.53 years took part in the experiment.
Out of the 15 participants, 13 were right-handed and seven were male. Seven of the participants had corrected vision. All participants were native speakers of Dutch.
After the experiment, each participant filled out a questionnaire about his or her knowledge of German.
According to this data, the participants on average had 12.69 years of experience with the German language with a standard deviation of 8.26 years. The participants seemed to have understood this question quite differently, while some gave the number of years since they first came in contact with German, others only stated the number of years during which they had a formal education in German.
Only 20 percent of participants declared that they read German regularly or often. On the other hand 53 percent stated that they hardly ever read texts in German.
Similarly, seven percent said they spoke German regularly or often, while 67 percent hardly ever speak German.
The participants were recruited by email through a participant list of the Nijmegen Institute for Cognition and Information and through posters hung up on the campus of Radboud University Nijmegen. Additionally, students on campus were approached and asked to take part in the experiment.
Participants were paid one Euro for every ten minutes that the experiment took.

## 3.4 Execution of the experiment

The experiment was executed in August 2005 at the University of Nijmegen. A computer running Apple OS 9 and a program written at the University of Nijmegen specifically for priming experiments were used.
Responses to the stimuli were given through a button box with three non-labeled horizontally aligned keys, only the left and right one of which were used. Right-handed participants were asked to press the right button if the string of letters presented was a word in Dutch or German and the left one otherwise. In order to avoid effects stemming from the handedness of the participant left-handed participants received the opposite instructions, that is, to press the button on the left if the item presented was a word and the right button if it was not.
The participants received oral instructions in German and written instructions in Dutch. They were told to answer as quickly and accurately as possible and to put the index finger of each hand on the corresponding button in order to be able to respond quickly and without a delay.
The instructions during the experiment were printed on the screen in Dutch. Each of the participants was presented with 416 items in total, a training set containing 16 items and four blocks of 100 items each. After each of the blocks the participants were encouraged to take a short break.
The training set consisted of eight non-words and eight Dutch and German words. While the training set was the same for each participant, there were different lists of the actual items to be tested for each of the participants in order to avoid effects stemming from the order of the items.
The instructions and items were presented in a white box measuring 450 by 250 pixels centered in the middle of the screen that was black otherwise. The font for both was Arial, the instructions were displayed at a size of 15 pt and the items at a size of 24 pt.
Before each item a star was displayed as a fixation mark for 500 ms in the middle of the screen. The prime and target words were then presented for 2000 ms. If no button was pressed before the deadline, the item was skipped and the next one was displayed. The inter-trial interval (ITI), a short pause between two consecutive trials, was set to 1000 milliseconds.

# 4 Results

The analysis of the experiment was made more difficult by the fact that the amount of knowledge of German differed a lot between the participants and for the most part was not very good.
While some of the participants were former or current students of German and thus were regularly exposed to German at a universitary level, most had hardly used the language since they completed secondary education.
This situation was not desired but could unfortunately not be avoided due to temporal restrictions in finding participants for the experiment.
The data were prepared for the analysis by first separating words from non-words. A significant non-word effect of over 100 ms was found in the data, both for the case that all responses ($t = 13.2174, p < 0.001$) and only correct responses ($t = 19.9452, p - value < 0.001$) were considered.
All incorrect responses were removed from the data.
The average response time and the standard deviation were computed and in all instances where the response to at least one of the items of a word pair lay 2.5 times the standard deviation above or under the average were removed.
An overview over the individual participants and items was constructed. All instances of items and their partners where the percentage of correct responses relative to the number of presentations was lower than the overall average number of correct responses minus 2.5 standard deviations were removed from the data.
The same was to be done for participants but because of the different degrees of knowledge of German the standard deviation was so high that this would have meant leaving in the data from subjects who performed hardly at chance level in German. The average rate of correct answers for German items was 0.59 with a standard deviation of 0.16 while the values for Dutch items were 0.88 and 0.04 respectively. In the case of German words, the minimum percentage of correct responses was only 32.32 percent while the maximum was 80.83 percent.
For the Dutch words the participant performing the best answered correctly in 93.55 percent of the cases while the smallest percentage of correct answers was 77.65 percent.
Because of these findings the data from six of the fifteen participants were excluded from further evaluation.
Since there are only two possible responses to each stimulus, fifty percent of correct answers is what participants would reach on average regardless of their knowledge of the language. As no participant performed at over 53 percent but under 66 percent for the German items, the data from the six

participants whose accuracy in responses to the German data was 53 percent or lower were removed.
The new average rate of correct responses and standard deviation for the German words were 0.72 and 0.06 with a new minimum value of 0.63.
This exclusion of data did not have a big effect on the values for the Dutch words, the average remained the same at 0.88 and the standard deviation slightly increased to 0.05.
As a consequence of removing the data of the six participants, both the average and the standard deviation of the reaction time for correct responses to Dutch items slightly decreased from 612.68 ms to 600.85 ms and from 192.33 ms to 168.00 ms respectively.
The effect on average reaction time and standard deviation for correct responses to German stimuli was stronger with values decreasing from 881.50 ms to 826.11 ms and 296.79 ms to 240.93 ms.
This was expected as performance in German was taken as a basis to make a decision about whether to exclude the data of a participant, and also because, as it is demonstrated in the frequency effect, more common items are not only recognized faster but also at a lower error rate. Because of this correlation, participants more familiar with German vocabulary do not only answer more correctly but also faster on average. Additionally, the amount of participants' knowledge of German was more heterogeneous than that of Dutch as can be seen from the initial standard deviations of accuracy and response time. This of course can be explained by the fact that all participants had Dutch as their native language and could thus be assumed to be roughly equally familiar with the Dutch items. Even after the removal a highly significant effect of language on reaction time ($t = 18.0663, p - value < 0.001$) could be found.
In the further analysis the prime and target instances of the individual German and Dutch words were averaged and matched.
This was done so that for each word a baseline for reaction time free of the language effect and other possible effects could be established through the average response time for the prime which itself of course if not primed. This way, any significant difference in reaction time can be attributed to the priming effect.

| | German | | Dutch | |
|---|---|---|---|---|
| +o-s-m | B | 750.68 | B | 618.28 |
| | P | 825.28 | P | 602.1162 |
| | D | -74.60 | D | 16.16 |
| +o-s+m | B | 815.24 | B | 556.64 |
| | P | 724.46 | P | 662.33 |
| | D | *90.78 | D | *-105.70 |
| -o+s-m | B | 855.93 | B | 579.51 |
| | P | 860.99 | P | 586.56 |
| | D | -5.06 | D | -7.05 |
| -o+s+m | B | 904.62 | B | 616.26 |
| | P | 859.83 | P | 606.21 |
| | D | 44.79 | D | 10.05 |
| -o-s+m | B | 793.03 | B | 636.02 |
| | P | 897.83 | P | 556.30 |
| | D | -104.80 | D | 79.72 |
| +o+s+m | B | 816.96 | B | 572.49 |
| | P | 771.01 | P | 585.6 |
| | D | 45.95 | D | -13.30 |
| +o+s-m | B | 784.15 | B | 598.20 |
| | P | 1051.92 | P | 564.35 |
| | D | *-267.78 | D | 33.84 |

Table 2: Average reaction times per item for the **B**ase and **P**rimed condition and their **D**ifference, significant priming effects are marked with a "*"

A Repeated-Measures ANOVA (Analysis of Variance) was executed.[11] ANOVA is similar to t-tests in that the outcome is the ratio between the variance between groups and the variation of scores within the sample groups. The difference between ANOVA and t-tests however is that the t-test only treats samples from two groups at once. ANOVA on the other hand can examine the differences in means for multiple factors and levels at once and then decides whether the differences in samples can be considered chance or whether they represent actual differences between the levels.
In terms of the t-test, Repeated-Measures ANOVAs can be compared to dependent samples t-tests. They take account of the fact that each participant does not only contribute one piece of data for one condition, but is tested

[11] For a basic introduction to ANOVA and the statistical concepts used in the following see [20][5][7]

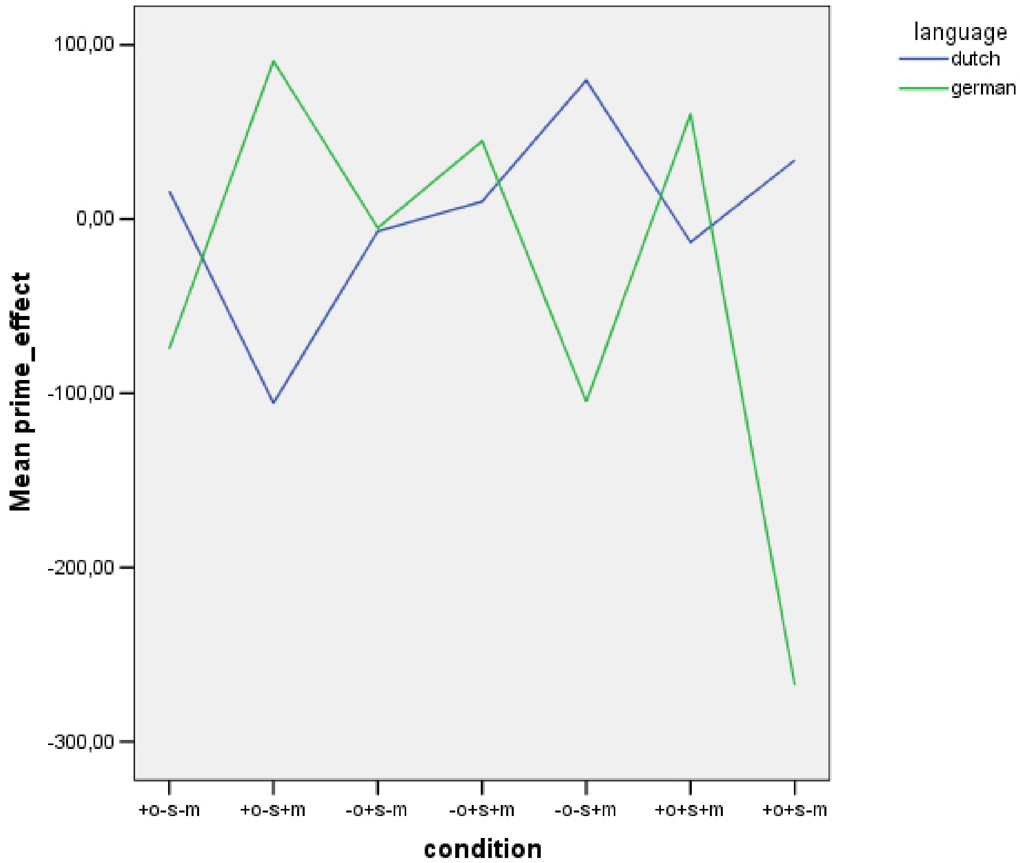


Figure 3: Average priming effects per category and language

several times with conditions varying.
The differences is that levels of factors are varied not between participants but for each participant ("within-subject") who thus contributes data to all levels. The advantage of this is that the individual variance between participants is known and can thus be taken into account in the analysis. Since the analysis of individual differences between participants is not relevant for the ANOVA, this means that the accuracy of the analysis is increased because parts of variation can be attributed to individual variation between the participants and thus neglected.
While Repeated-Measure ANOVAs require generally less participants due to the increased accuracy, they cannot be used if the residual degrees of freedom are too low.
Unfortunately, given the small amount of remaining data, and the high number of factors and levels which totaled 28 different between-subject levels, the ANOVA did not return any results due to a lack of data.
This outcome of course also prohibited a detailed analysis by lag which might have been provided further insights into the time course of the different priming effects.
Since the high number of levels given the relatively small amount of data turned out to be so problematic, the number of levels was reduced by merging the seven different configurations into three. This was done by combining all configurations that contained "+o" into the first new level, all that contained "+s" into the second and all that contained "+m" into the third. For example, the reaction times to "+o-s-m" word pairs, as well as those to "+o-s+m", "+o+s+m" and "+o+s-m" word pairs were all included in the new level "+o".
A Repeated-Means ANOVA with language (Dutch and German), primed (baseline and primed) and combined categories ("+o", "+s" and "+m") as the factors was run on the input.
Significant effects of combined category ($p = 0.04$), language ($p < 0.001$) and the combination of combined category and language ($p = 0.01$) could be detected.
The combination of combined category and primed approached significance at $p = 0.15$ but did not become significant. The reason for this (apart from the obvious reason that there might be no priming effect) might be the big standard deviations in the response times of the different participants which lessens the significance of the effect and which in turn can be attributed to the low number of participants.
However the problem of not being able to analyze the data for the individual categories due to a lack of data remains. As mentioned before both t-tests and ANOVA test whether the difference in means between groups tests could

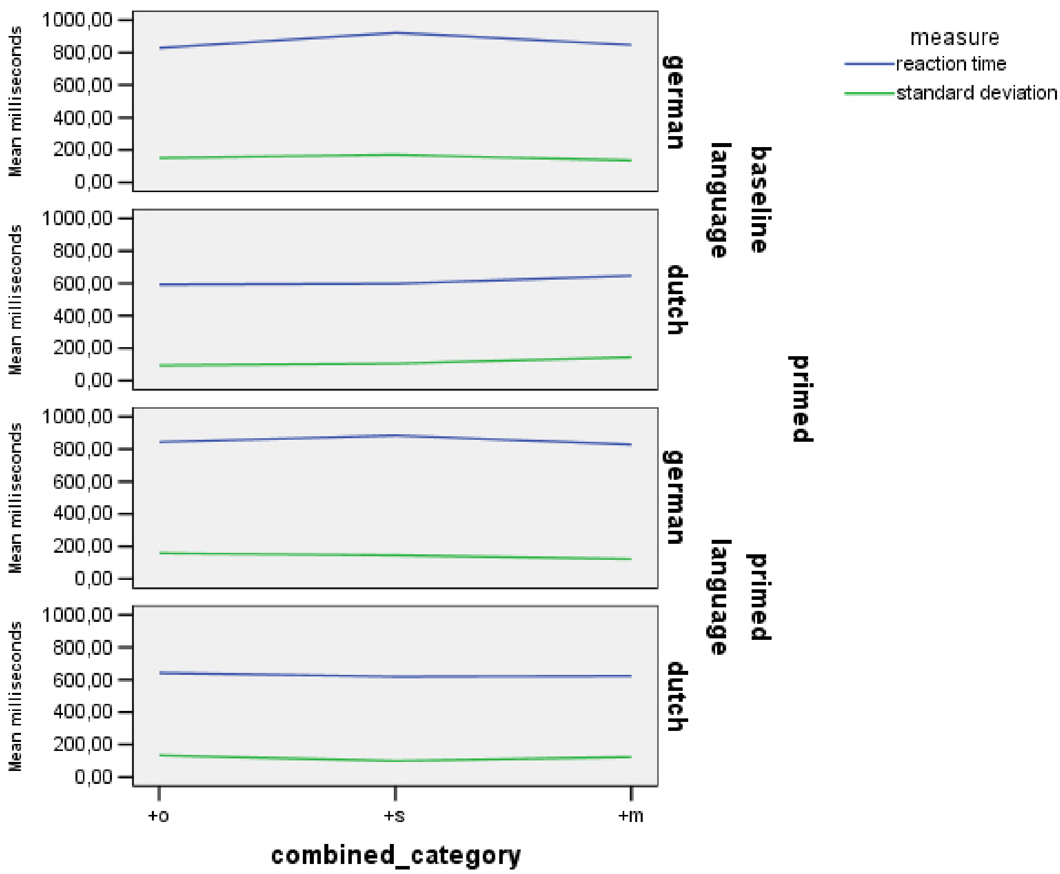


Figure 4: Average reaction times and standard deviations per combined category

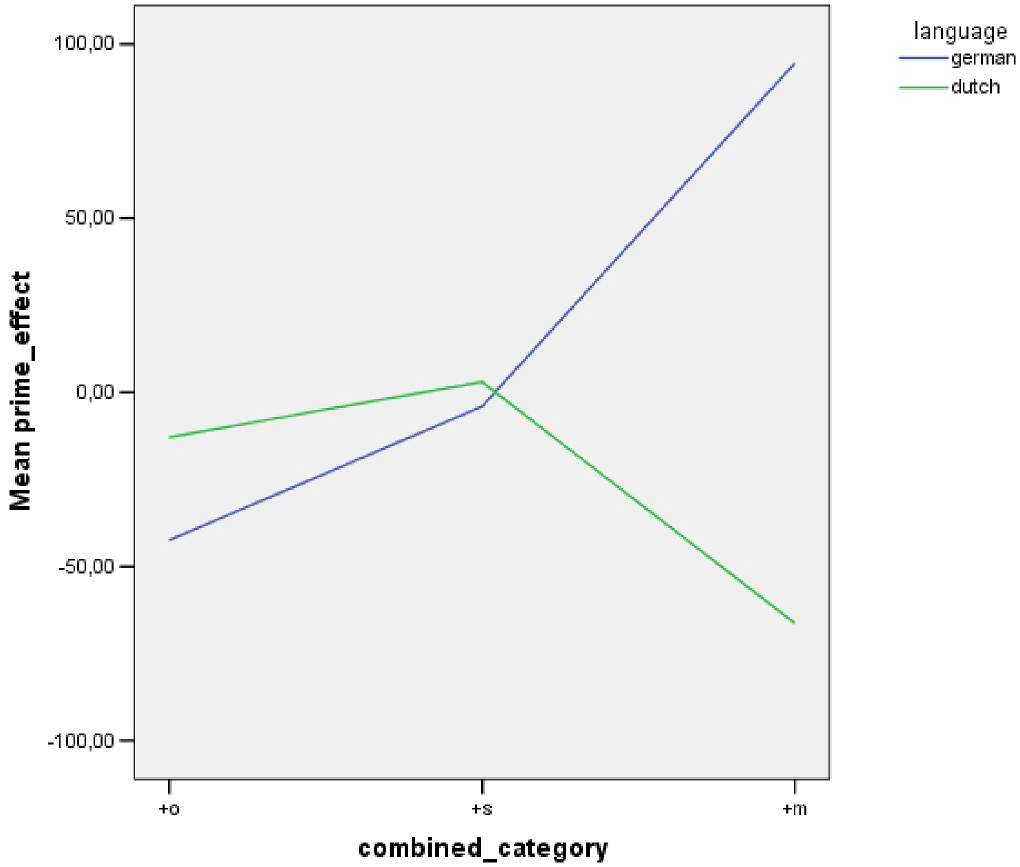


Figure 5: Average priming effects per combined category and language

be caused by a sampling error and thus have similar applications.
The basic difference in terms of application between t-tests and ANOVA is that the latter is more general in that it cannot only test hypotheses about two population means, but analyze several groups in one go and thus detect interaction effects that can not be found using t-tests.
In principle it is possible to do pairwise t-tests instead of an ANOVA, despite the lesser power of t-tests. The problem with this (apart from the fact that it is far more inconvenient) is that the alpha error, the probability of finding a significant difference when it does not actually exist, increases with the number of tests executed using the same data. The formula for this so-called family-wise error rate is $1-(1-\alpha)^k$ where k is the number of tests conducted on the same data.
So for example using t-tests to find out whether the individual priming effects per language and category are significant poses no problem in terms of an increased error rate since each set of data (for example the reaction times to a prime, that is, the baseline cases for German items in the "+o+s+m" condition) is only used once.
However, if the significance of the priming effects between categories is tested, each set of data is used six times since it must be compared to the data from all other six categories. This means that, given $\alpha = 0.05$, that the family-wise error rate increases to $1 - (1 - 0.05)^6 = 0.26$ which means that there is a 26 percent chance of an $\alpha$-error.
A way to improve on this high chance of an error is to set the significance level $\alpha$ to a lower value, that is, to require a smaller probability of an $\alpha$-error in order to consider the result of the t-test significant and reject the null hypothesis that the means of the two samples are equal.
If $\alpha$ is set to 0.02, the family-wise error rate in our case is $1-(1-0.02)^6 = 0.11$. The probability of a type I error thus is less than half of what it would be for $\alpha = 0.05$.
For this reason the significance of single priming effects within categories and languages was tested using t-tests and an $\alpha$ of 0.05 and inter-category priming effects were tested using a significance level of 0.02.
It must be stressed that this procedure is not equal in power and accuracy to the intended ANOVA. However, given the small amount of data it offers a relatively reliable and informative way to analyze the data.
In German only the priming effects of "-o-s+m" ($t = -1.7826, df = 18.704, p-value = 0.050$) and "+o+s-m" ($t = -2.4974, p - value = 0.03496$) were significant with "+o-s+m" approaching significance at $p = 0.081$ ($t = 1.826, p-value = 0.081$).
In the Dutch data only the priming effect in "+o-s+m" was significant ($t = -2.0869, p - value = 0.049$).

The differences between priming effects were significant for categories "+o-s+m" and "-o-s+m" ($t = -2.5654, p - value = 0.01784$) in Dutch and categories "+o-s+m" and "-o-s+m" ($t = 2.9071, p-value = 0.008$) and categories "+o+s+m" and "+o+s-m" ($t = 2.92, p - value = 0.01843$) in German (see table).

| German | | Dutch | |
|---|---|---|---|
| -o-s+m | t = -1.7826<br>p-value = 0.050 | +o-s+m | t = -2.0869<br>p-value = 0.049 |
| +o+s-m | t = -2.4974<br>p-value = 0.03496 | | |

Table 3: Significant priming effects

| German | | Dutch | |
|---|---|---|---|
| +o-s+m<br>and -o-s+m | t = 2.9071<br>p-value = 0.008 | +o-s+m<br>and -o-s+m | t = -2.5654<br>p-value = 0.01784 |
| +o+s+m<br>and +o+s-m | t = 2.92<br>p-value = 0.01843 | | |

Table 4: Significant differences between priming effects

# 5 Discussion

As can be seen from Table 2, priming with words that were only semantically related had practically no effect on reaction times. As mentioned, the results on semantic priming are divided between studies that find priming effects for translation equivalents and those that do not and claim that semantic priming effects can only be found when using a task that requires more semantic processing of the items (like animacy decision where participants have to categorize each word as referring to an animate or inanimate object). Priming with orthographically similar items had a negative priming effect which is in line with both the findings from crosslinguistic as well as monolinguistic findings in similar studies.

Priming with only morphologically related words had a negative effect on reaction times on German items but a positive one on the reaction time for Dutch words.

The reaction times in the table also illustrate the fact that reaction times for German items in all cases are at least 100 millseconds longer than those to Dutch items.

Another thing that is evident from the overview over reaction times is that priming effects are generally much stronger for German items. That is, if the prime is Dutch and the target German, the reaction time to the prime differs more from the baseline reaction time for the item than when the prime is German and the target Dutch.

This could be explained by the fact that, although participants performing too badly on German items were removed, the amount of correct answers where the participant just happened to guess correctly without really recognizing the word can be assumed to be higher for German items than for Dutch items. This would mean that the Dutch targets are not actually primed (due to a failure to recognize the German prime), which is why there is no big difference between the baseline reaction time and the primed reaction time. However, for this explanation to be plausible, it would also have to be assumed that priming with a Dutch word increases the probability to recognize the corresponding German target.

Addionally, at least purely orthographic priming should then not display this difference in the amount of priming between Dutch and German since no semantic or morphological analysis and activation is needed and orthographic similarity can be recognized without knowing the meaning or structure of the word. But as can be seen from the table, this prediction is not consistent with the actual data.

Another possible explanation is that this effect does not stem from the participants' failure to recognize the German words but that from the fact that

German is on their second or even third language which many of them do not use very often. Because of this the recognition is less unique and reliable which formulated in terms of activation means that the activation upon presentation of a German word is lower and more spread out, causing a weaker priming effect.
One striking effect that is especially noticeable in the graphic is that the priming effect on Dutch items in all conditions mirrors the priming effect for German items in the same category. That is, in the conditions where there is a positive priming effect for German, there is a negative priming effect on Dutch items and vice versa. This effect was not predicted by previous studies and might be interesting for further investigation.
While the priming effects for the purely semantic, orthographical and morphological conditions are as expected, none of them is significant which could be attributed to the small number of participants and the high standard deviations mentioned before.
Also, the case of conditions where several levels of similarity were combined were less clear. Priming with cognates (thus "+o+s+m") as expected and as predicted from other studies had a facilitatory effect in the case that the prime was Dutch and the target was German.
The priming effects for the combined categories, that is, those, that contained more than one element with a "+"-value, were clearly not simply additive. For example the inhibitory effect of purely orthographic priming on German targets and the slightly weaker facilitatory effect of purely morphological priming on German targets would be excepted to add up to a weak inhibitory priming effect for the "+o-s+m"-condition in German if the priming effects were simply added up equally.
The priming effects of the single components become more clear in the analysis that employed combined categories.
Since all conditions that contain a feature are combined in the corresponding combined category, the combined priming effect can help discerning the general nature of the different priming effects of "+o", "+s" and "+m", taking into consideration not only the conditions with just one similarity but also the ones with combinations of active features.
The combined categories for orthographic and semantic similarity show similar effects as the purely orthographic and semantic categories in the by-condition analysis, namely an -albeit weaker- negative priming effect from orthography and next to no priming effect from semantics in German and only very weak priming effects for both in Dutch.
The nature of the morphological effect is less clear, though. While the purely morphologically related primes have a positive priming effect on Dutch targets, they have a negative effect on German targets.

On the other hand, the combined morphological category shows the reverse effects, namely a facilitatory effect on German targets and an inhibitory one on Dutch targets which stems from the fact that all conditions that contain a morphological component along with at least another component have a positive priming effect in German and a negative effect in Dutch.
Since only two of the seven conditions show significant priming effects in German and only one does in Dutch, this could in theory be attributed to chance. But the fact that the observation holds without exception both for German and Dutch and given that the priming effects for orthography, semantics and cognates are in accordance with previous findings, suggests that the observation should not be dismissed as easily.
As for the other main question of this study, the thesis that there is a distinctively morphological component in cognate priming could be verified through the finding that the priming effect for "+o+s+m" (cognates) and "+o+s-m" was found to be significantly different.
This means that cognate recognition is not only guided by similar orthographic form and meaning, but that there are also morphological processes involved.
This significance could only be established for German targets as priming effects on Dutch targets were generally much weaker.

## 5.1 Conclusion

The main question of this study could be answered successfully: There is a statistically significant difference between the priming effects for cognates and that for only orthographically and semantically similar words from which can be concluded that there is a morphological component in cognate priming.
Additionally, the results indicate that results for orthographical and semantic crosslinguistic priming from previous studies could be replicated. The fact that these two effects are similar in crosslinguistic as well as in studies employing only one languages provides further evidence for the concept of an integrated lexicon and processes in word recognition.
Another finding is that the priming effects from Dutch primes on German targets were significantly stronger than the priming effect from German primes on Dutch targets and that, while they are weaker, the Dutch priming effects in many cases mirror the German priming effects.
But the statistical analysis of these findings was only in part successful, requiring further investigation. So while it could be shown that there is a morphological component to cognate priming, the other findings in part should

be considered rather as pointers than as definite results and should be further investigated in a follow-up study employing more participants with generally a better and more consistent knowledge of German. This also would make it possible to study the time course of the different priming effects. Since this would mean that there are four factors (language, primed/not primed, lag and condition) with a total of fifty-six levels, the number of participants would have to be far bigger than the nine that this study used. This might also help avoid the problem of high standard deviations in reaction times which posed a problem for the analysis of this experiment.
Further investigation might also explore the interrelation between native language (Dutch or German) and proficiency in the other language (German or Dutch) and the effect on performance in the experimental setup.
And as the main question of this study has been answered successfully, all these further investigation possibilities seem very promising for the prospect of finding out more about the role of crosslinguistic morphological priming.

# 6 Appendix

| Condition | Items |
|---|---|
| +o+s+m | BESTEMMING-BESTIMMUNG |
| | DOELLOOS-ZIELLOS |
| | DOMHEID-DUMMHEIT |
| | GODHEID-GOTTHEIT |
| | HONGERIG-HUNGRIG |
| | KAMERAADSCHAP-KAMERADSCHAFT |
| | KRIJGER-KRIEGER |
| | LEVENDIG-LEBENDIG |
| | LIEFLIJK-LIEBLICH |
| | LUCHTIG-LUFTIG |
| | MEESTERLIJK-MEISTERLICH |
| | NAAKTHEID-NACKTHEIT |
| | VERLIEZER-VERLIERER |
| | VETTIG-FETTIG |
| | VRAGER-FRAGER |
| | WAAKZAAM-WACHSAM |
| | WONDERLIJK-WUNDERLICH |
| +o+s-m | BOUWER-BAUHERR |
| | DIERLIJK-TIERISCH |
| | DRONKENSCHAP-TRUNKENHEIT |
| | ERFENIS-ERBTEIL |
| | FEITELIJK-FAKTISCH |
| | KUNSTMATIG-KUENSTLICH |
| | REGENACHTIG-REGNERISCH |
| | PARADIJSELIJK-PARADIESISCH |
| +o-s+m | BEZIGHEID-EWIGKEIT |
| | HEVIGHEID-HEFTIGKEIT |
| | JARIG-FAHRIG |
| | LETTERLIJK-ELTERLICH |
| | MATIG-MUTIG |
| | NAAMLOOS-NAHTLOS |
| | OUDERLIJK-SONDERLICH |
| | ROESTIG-POETISCH |
| | RUMOERIG-NUMERISCH |
| | SNELHEID-NEUHEIT |

| | |
|---|---|
| | TIENER-EINER |
| | VREDIG-FREUDIG |
| +o-s-m | BOKSER-BOHRUNG |
| | DEKSEL-DEUTUNG |
| | DRUKTE-DREHUNG |
| | EENHEID-AHNUNG |
| | GEZINDHEID-FEINDSCHAFT |
| | KLEVERIG-KLEINLICH |
| | LULLIG-LUSTLOS |
| | MENGSEL-MELKER |
| | REDDING-REINHEIT |
| | SPEEKSEL-SPENDER |
| | VEILING-FEIGHEIT |
| | VESTIGING-FERTIGKEIT |
| -o+s+m | AVONTUURLIJK-ABENTEUERLICH |
| | BIJZONDERHEID-BESONDERHEIT |
| | DROEVIG-TRAURIG |
| | GELIJKHEID-GLEICHHEIT |
| | HAVELOOS-MITTELLOS |
| | HUMEURIG-LAUNISCH |
| | IJZEREN-EISERN |
| | KLEURIG-FARBIG |
| | LAFHEID-FEIGHEIT |
| | SCHOOIER-BETTLER |
| | TIJDIG-ZEITIG |
| | TIJDLOOS-ZEITLOS |
| | VENNOOTSCHAP-PARTNERSCHAFT |
| | VRIJER-FREIER |
| | WERKZAAM-ARBEITSAM |
| | WIJZER-ZEIGER |
| | ZEDELIJK-SITTLICH |
| | ZEKERHEID-SICHERHEIT |
| -o+s-m | ARMOEDIG-AERMLICH |
| | BEGINSEL-GRUNDSATZ |
| | BEKOORLIJK-ANZIEHEND |
| | DUISTERNIS-DUNKELHEIT |
| | GEBREKKIG-MANGELHAFT |

| | |
|---|---|
| | LAFAARD-FEIGLING |
| | RILLING-SCHAUDER |
| | SCHAARSTE-KNAPPHEIT |
| | SCHEMERING-ZWIELICHT |
| | SCHEPSEL-KREATUR |
| | STAKING-AUSSTAND |
| | VIJANDIG-FEINDLICH |
| | VOEDING-MAHLZEIT |
| | VOEDZAAM-NAHRHAFT |
| | WISSELVALLIG-WECHSELHAFT |
| | ZIEKELIJK-KRANKHAFT |
| -o-s+m | BALLINGSCHAP-BRUDERSCHAFT |
| | BROMMER-DAMPFER |
| | FIGUURLIJK-STERBLICH |
| | GEBRUIKER-LAUSCHER |
| | GELDIGHEID-TAPFERKEIT |
| | GELOVIG-BLUTIG |
| | HEERSER-KRATZER |
| | HOUDER-FECHTER |
| | KIJKER-MALER |
| | KWETSBAAR-BRAUCHBAR |
| | ONDERWIJZER-SCHLEICHER |
| | TEGENWOORDIGHEID-GESCHLOSSENHEIT |
| | VROLIJKHEID-SAUBERKEIT |
| | WAARDIG-WINDIG |
| | WERKLOOS-FURCHTLOS |
| | ZWIJGZAAM-EMPFINDSAM |

Table 5: The stimulus material used in the experiment (without the items that were discarded due too few correct answers)

# References

[1] Shlomo Bentin and Laurie Feldman. The contribution of morphological and semantic relatedness to repetition priming at short and long lags: Evidence from hebrew. *The Quarterly Journal of Experimental Psychology*, 42A(4):693–711, 1990.

[2] Jean Berko. The child's learning of english morphology. *Word*, 14:150–177, 1958.

[3] Ranka Bijeljac-Babic, Agnes Biardeau, and Jonathan Grainger. Masked orthographic priming in bilingual word recognition. *Memory and Cognition*, 25(4):447–457, 1997.

[4] Geert Booij. Inflection and derivation. In Keith Brown, editor, *Encyclopedia of Language and Linguistics (2nd edition)*. Oxford: Elsevier, 2005.

[5] Scott Brown. Spss guide part 2. http://oz.ss.uci.edu/PDFS/spssman2.pdf.

[6] Sean Crist. Toward a formal markup standard for etymological data, 2005.

[7] Jamie DeCoster. Testing group differences using t-tests, anova, and nonparametric measures. http://www.stat-help.com/ANOVA.pdf.

[8] Kevin Diependaele, Dominiek Sandra, and Jonathan Grainger. Masked cross-modal morphological priming: Unravelling morpho-orthographic and morpho-semantic influences in early word recognition. *Language and Cognitive Processes*, 20(1–2):75–714, 2005.

[9] Ton Dijkstra, Fermì Moscoso del Prado Martìn, Béryl Schulpen, Robert Schreuder, and Harald Baayen. A roommate in cream: Morphological family size effects on interlingual homograph recognition. *Language and Cognitive Processes*, 20(1–2):7–41, 2005.

[10] Ton Dijkstra, Mark Timmermans, and Herbert Schriefers. On being blinded by your other language: Effects of task demands on interlingual homograph recognition. *Journal of Memory and Language*, 42(4):445–464, 2000.

[11] Alberto Dominguez, Juan Segui, and Fernando Cuetos. Morphological processing in word recognition: A review with particular reference to spanish data. *Psicològica*, 21(2):375–401, 2000.

[12] Etta Drews and Pienie Zwisterlood. Morphological and orthographic similarity in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 21(5):1098–1116, 1995.

[13] Wouter Duyck, Kevin Diependale, Denic Drieghe, and Marc Brysbaert. The size of the cross-lingual masked phonological priming effect does not depend on second language proficiency. *Experimental Psychology*, 51(2):1–9, 2004.

[14] Laurie Feldman and Jasmina Moskovljevic. Repetition priming is not purely episodic in origin. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(4):573–581, 1987.

[15] Laurie Feldman and Brendon Prostko. Graded aspects of morphological processing: Task and processing time. *Brain and Language*, 81:12–27, 2002.

[16] Kenneth Forster and Chris Davis. Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(4):680–698, 1984.

[17] Carol Fowler, Shirley Napps, and Laurie Feldman. Relations among regular and irregular morphologically related words in the lexicon as revealed by repetition priming. *Memory and Cognition*, 13(3):241–255, 1985.

[18] Robert Frick. Accepting the null hypothesis. *Memory and Cognition*, 23(1):132–138, 1995.

[19] Ram Frost, Avital Deutsch, Orna Gilboa, Michal Tannenbaum, and William Marslen-Wilson. Morphological priming: Dissociation of phonological, semantic, and morphological factors. *Memory and Cognition*, 28(8):1277–1288, 2000.

[20] Burt Gerstman. Statprimer (version 6.2). http://www2.sjsu.edu/faculty/gerstman/StatPrimer/.

[21] Hélène Giraudo and Jonathan Grainger. Effects of prime word frequency and cumulative root frequency in masked morphological priming. *Language and Cognitive Processes*, 15(4–5):421–444, 2000.

[22] Laura Gonnerman and Mark Seidenberg. Explaining derivational morphology as the convergence of codes. *Trends in Cognitive Sciences*, 4(9):353–361, 2000.

[23] Laura Gonnerman, Mark Seidenberg, and Elaine Andersen. Graded semantic and phonological similarity effects in priming: Evidence for a distributed connectionist approach to morphology.

[24] Jonathan Grainger and Cheryl Frenck-Mestre. Masked priming by translation equivalents in proficient bilinguals. *Language and Cognitive Processes*, 13(6):601–623, 1998.

[25] Vicki Hansen and Laurie Feldman. Language specificity in lexical organization: Evidence from deaf signers' lexical organization of asl and english. *Memory and Cognition*, 17(3):292–301, 1989.

[26] Juhani Jaervikivi and Jussi Niemi. Form-based representation in the mental lexicon: Priming (with) bound stem allomorphs in finnish. *Brain and Language*, 81(1–3):412–423, 2002.

[27] Daniel Jurafsky and James Martin. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition.* Prentice Hall PTR, 2000.

[28] Catharine Keatley, John Spinks, and Beatrice De Gelder. Asymmetrical cross-language priming effects. *Memory and Cognition*, 22(1):70–84, 1994.

[29] Catherine-Marie Longtin, Juan Segui, and Pierre Hallé. Morphological priming without morphological relationship. *Language and Cognitive Processes*, 18(3):313–334, 2003.

[30] William Marslen-Wilson, Lorraine Komisarjevsky Tyler, Rachelle Waksler, and Lianne Older. Morphology and meaning in the english mental lexicon. *Psychological Review*, 101(1):3–33, 1994.

[31] James McQueen and Anne Cutler. Morphology in word recognition. In Andrew Spencer and Arnold Zwicky, editors, *Handbook of Morphology.* Oxford: Blackwell, 1998.

[32] Fanny Meunier and Juan Segui. Cross-modal morphological priming in french. *Brain and Language*, 81(1–3):89–102, 2002.

[33] Stephen Monsell. Repetition and the lexicon. In Andrew Ellis, editor, *Progress in the Psychology of Language*, volume 2, pages 147–195. Erlbaum, London, 1985.

[34] James Neely. Semantic priming effects in visual word recognition: A selective review of current findings and theories. In Derek Besner and Glyn Humphreys, editors, *Basic processes in reading: Visual word recognition*. Hillsdale, NJ: Erlbaum, 1991.

[35] Scott Piao, Paul Clough, and John Arundel. Proposing basic approaches to detecting text rewrite, 2001.

[36] Kathy Rastle, Matt Davis, William Marslen-Wilson, and Lorraine Tyler. Morphological and semantic effects in visual word recognition: A time-course study. *Language and Cognitive Processes*, 15(4–5):507–537, 2000.

[37] Jennifer Stolz and Laurie Feldman. The role of orthographic and semantic transparency of the base morpheme in morphological processing. In Laurie Feldman, editor, *Morphological aspects of language processing*, pages 109–129. Erlbaum, Hillsdale, NJ, 1995.

[38] Marcus Taft and Kenneth Forster. Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 14(6):638–647, 1975.

[39] Michael Thomas and Kim Plunkett. Representing the bilingual's two lexicons. In *COGNITIVE SCIENCE SOCIETY: 17TH ANNUAL CONFERENCE*, 1995.

[40] Walter Van Heuven, Ton Dijkstra, and Jonathan Grainger. On being blinded by your other language: Effects of task demands on interlingual homograph recognition. *Journal of Memory and Language*, 39(3):458–483, 1998.

[41] René Zeelenberg and Diane Pecher. False memories and lexical decision: even twelve primes do not cause long-term semantic priming. *Acta Psychologica*, 15(3):269–284, 2002.

[42] René Zeelenberg and Diane Pecher. Evidence for long-term cross-language repetition priming in conceptual implicit memory tasks. *Journal of Memory and Language*, 49(1):80–94, 2003.

[43] Johannes Ziegler, Conrad Perry, Anna Ma-Wyatt, Diana Ladner, and Schule-Koerne Gerd. Developmental dyslexia in different languages: Language-specific or universal? *Journal of Experimental Child Psychology*, 86(3):169–193, 2003.

I hereby confirm that I wrote this thesis independently and that I have not made use of any other resources or means than those indicated.

Hiermit bestätige ich, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Düsseldorf, 21.11.2005