Congruency and typicality effects in lexical decision

Sebastian Loth

Royal Holloway, University of London

Ph.D.

Declaration of Authorship

I Sebastian Loth hereby declare that this thesis and the work presented in it is entirely my own. Where I have consulted the work of others, this is always clearly stated.

Acknowledgements

I would like to thank my supervisor Colin Davis for giving me the opportunity to start and the support and the expert advice to complete this PhD, but also for being patient, encouraging and a friend all the way. I would like to thank all the people who have contributed to this work through discussions and advice, especially Sally Andrews, Dave Balota, Mark Brysbaert, Polly Dalton, Thomas Pechmann and Kathy Rastle.

I thank Luise Dorenbusch. I am dearly grateful that I could always rely on you and your support.

I would like to thank Fiona Buxton and her staff for their care and commitment.

I thank Leonie Brose, the Zankers and the Dorenbuschs for their support and providing shelter. For their support and company, I would like to thank Linda Bayliss, Carol Blackman, Davide Crepaldi, Lise van der Haegen, Ana de Macedo, Sam McCormick, Marjolein Merkx, Lesley Mitchell, Lissy Otto and Helena Trompelt. Especially, I would like to thank Johannes Zanker for making Royal Holloway the friendly place that I will remember.

I would like to thank my granny and my mum for their continuous help and support.

Abstract

This thesis describes basic research into visual word recognition and decision making. Determining the best matching lexical representation for a given stimulus involves interactions between representations. The standard task for studying these processes is the lexical decision task (LDT), but there is still debate regarding the factors that affect how individuals make lexical decisions. The nature of lexical interactions and the processes underlying lexical decision-making were addressed here by testing response congruency effects in the masked priming variant of the LDT.

The results of seven masked priming experiments showed a robust response congruency effect that depends on the difficulty of the word-nonword discrimination. This finding resolved apparent inconsistencies in previous research. The experiments were simulated using the Bayesian Reader and the Spatial Coding Model (SCM). The probability based Bayesian Reader model failed to accommodate the findings. However, a good fit to the data was provided by a modified version of the SCM in which the assumptions regarding the nature of lexical interactions were changed such that word nodes inhibit only (closely) related competitors. The model also assumes that the difficulty of the word-nonword discrimination affects the degree to which stimulus typicality informs lexical decisions.

A critical issue for these experiments involved the definition of orthographic typicality. An algorithm for measuring orthographic typicality and for generating nonwords with a specific level of orthographic typicality (OT3) was developed. An unprimed LDT experiment showed that OT3 affected decision latency even when other standard measures of orthographic typicality were controlled. Two additional masked priming experiments showed that highly typical primes lead to faster word responses and slower nonword responses than less typical primes. Overall, the results of this research enhance our understanding of the processes underlying visual word recognition and lexical decision making, and also have important methodological implications for the field.

Table of Contents

Declaration of Authorship....2
Acknowledgements....3
Abstract....4
0. Preface....14
1. Response congruency effects....17
1.1 Masked priming....17
1.1.1 Effective masking....18
1.1.2 Strategic effects....19
1.1.3 Masked versus unmasked priming....20
1.1.4 Prime duration....21
1.1.5 Conclusion....22
1.2 Theoretical accounts of response congruency effects....23
1.2.1 Depth of processing....23
1.2.1.1 Deep Processing....23
1.2.1.2 Stimulus-response mapping....24
1.2.1.3 Prime novelty....25
1.2.1.4 Summary....25
1.2.2 Processes influencing priming effects....26
1.2.2.1 Action triggers....26
1.2.2.2 Two processes model....27
1.2.2.3 Conclusion....28
1.2.3 Summary....28
1.3 Properties of the stimuli and the experiment....29
1.3.1 Category size....29
1.3.1.1 Cardinality....29
1.3.1.2 Category coherence....30
1.3.1.3 Impostors....32
1.3.2 Task expertise....33
1.3.3 Response speed....33
1.3.4 Conclusion....35
1.4 Response congruency effects in lexical decision....35
1.4.1 Predictions from theoretical accounts....36
1.4.1.1 Deep processing and stimulus-response mapping....36
1.4.1.2 Category size....37
1.4.2 Empirical evidence....38
1.4.2.1 Studies reporting null effects....38
1.4.2.2 Studies reporting an effect....38
1.4.2.3 Empirical discrepancy....40
1.4.3 Conclusions....41
1.5 Summary....41
2. Orthographic typicality....43
2.1 Review of orthographic typicality....43
2.1.1 Lexical frequency....43
2.1.2 Neighbours....44
Relative frequency of neighbours....46
Other types of neighbours....47

2.1.3 Global measures of typicality....49
Bigrams in tachistoscopic identification....50
Bigram frequency in lexical decision....53
2.2 Algorithms for generating nonwords....57
2.2.1 WordGen....58
2.2.2 LINGUA....60
2.2.3 Wuggy....61
2.2.4 Summary....63
2.3 Orthographic trigram transition typicality (OT3)....63
2.3.1 Differences to other algorithms....64
Frequency and transition probabilities....64
Position and length dependent measures....66
Bigrams and trigrams....67
Sum and product....70
2.3.2 OT3 Algorithm....71
2.3.3 Initialising....72
2.3.4 Analysing words....75
2.3.5 Generating items....78
2.3.6 Summary....81
2.4 Experiment 1....81
2.4.1 Methods....81
2.4.2 Results....84
Repeated measures analysis of variance....84
Analysis in deciles....86
Ex-Gaussian analysis....87
2.4.3 Discussion....89
2.5 Summary and Discussion....91
2.6 Conclusion....92
3. Experiment 2....94
3.1 Introduction....94
3.2 Methods....95
3.3 Results....97
3.3.1 Repeated measures analysis of variance....98
3.3.2 Analysis in deciles....99
3.3.3 Analysis using the Ex-Gaussian fit....99
3.4 Discussion....100
3.4.1 Empirical discrepancy....100
Prime novelty and adaptation....101
Difficulty of the task....102
3.4.2 Theoretical accounts....102
Deep processing account....102
Semantic overlap model....103
3.5 Conclusion....103
4. Computational Models....105
4.1 Bayesian Reader....105
4.1.1 Overview....105
4.1.2 Simulation....107

4.1.3 Conclusion....110
4.2 Spatial Coding Model....111
4.2.1 Overview....111
4.2.2 Simulation....114
4.2.2.1 Nonword targets....114
4.2.2.2 Word targets....115
4.2.3 Inhibition in the Spatial Coding Model....116
4.2.3.1 Homogeneous inhibition....117
4.2.3.2 Selective inhibition....118
Overview....119
Connection weights....119
Computational load....120
4.2.4 Simulation in selective inhibition SCM....122
4.2.5 Summary....124
4.3 Benchmark simulation....125
4.3.1 Bayesian Reader....126
4.3.2 Selective inhibition Spatial Coding Model....126
4.4 Summary and Discussion....127
4.4.1 Homogeneous and selective inhibition....127
4.4.2 Target difficulty....128
Reduced effect....128
Decision models....129
Increased effect....131
5. Effects of target difficulty....132
5.1 Empirical discrepancy....132
5.2 Experiment 3....134
5.2.1 Methods....134
5.2.2 Results....135
Repeated measures analysis of variance....135
Analysis in deciles....136
Ex-Gaussian analysis....137
5.2.3 Discussion....137
5.3 Experiment 4....138
5.3.1 Methods....139
5.3.2 Results....139
Repeated measures analysis of variance....140
Analysis in deciles....141
Ex-Gaussian analysis....142
5.3.3 Discussion....142
5.4 Experiment 5....144
5.4.1 Methods....144
5.4.2 Results....144
Repeated measures analysis of variance....145
Analysis in deciles....146
Ex-Gaussian analysis....146
5.4.3 Discussion....147
5.5 General discussion....147

5.5.1 Distributional analysis....148
5.5.2 Negative congruency effects....149
5.5.3 Theoretical accounts of response congruency....152
5.5.4 Empirical discrepancy....153
5.6 Conclusion....153
6. Effects of prime difficulty....155
6.1 Prime informativeness....155
6.2 Experiment 6....155
6.2.1 Methods....156
6.2.2 Results....156
Repeated measures analysis of variance....157
Analysis in deciles....158
Ex-Gaussian analysis....158
6.2.3 Discussion....159
6.3 Experiment 7....159
6.3.1 Methods....160
6.3.2 Results....160
Repeated measures analysis of variance....161
Analysis in deciles....162
Ex-Gaussian....162
6.3.3 Discussion....163
6.4 Experiment 8....164
6.4.1 Methods....165
6.4.2 Results....165
Repeated measures analysis of variance....165
Analysis in deciles....166
Ex-Gaussian analysis....167
6.4.3 Discussion....167
6.5 Discussion....168
6.5.1 Word targets....169
6.5.2 Nonword targets....170
6.5.3 Prime informativeness....171
6.5.4 Theoretical accounts of response congruency effects....172
6.6 Summary....173
7. Effects of prime typicality....175
7.1 Experiment 9....177
7.1.1 Methods....177
7.1.2 Results....178
Repeated measures analysis of variance....178
Analysis in deciles....180
Ex-Gaussian analysis....180
7.1.3 Discussion....181
7.2 Experiment 10....182
7.2.1 Methods....183
7.2.2 Results....184
Repeated measures analysis of variance....184
Analysis in deciles....185

Ex-Gaussian analysis....186
7.2.3 Discussion....186
7.3 Discussion....187
Negative typicality effects....188
Theoretical accounts of response congruency....189
7.4 Conclusion....190
8. Simulations in the Spatial Coding Model....192
8.1 Verification in the Spatial Coding Model....192
8.2 Fitting the Spatial Coding Model....196
8.2.1 Empirical data....196
8.2.2 Parameters....197
8.2.3 Parameter search....199
8.3 Results of the simulations....203
8.4 Discussion....205
8.4.1 Typicality of nonwords....206
8.4.2 Task difficulty....207
8.4.3 Impact of replacing primes....209
8.4.4 Model modifications....211
8.5 Summary....212
9. General Issues....214
9.1 Summary empirical findings....214
9.1.1 Response congruency effect....214
9.1.2 Typicality effect....217
9.1.3 Negative typicality effects....218
9.1.4 Relation of typicality priming to other priming effects....220
9.1.5 Theoretical accounts of response congruency....222
9.2 Modelling response congruency and typicality effects....224
9.3 Measuring orthographic typicality....226
9.4 What is the purpose of selective inhibition?....227
9.5 What is the purpose of orthographic typicality?....229
References....231
Appendix A: Analysis of prime novelty and adaptation in Experiment 2....257
A.1 Prime novelty....257
Analysis....257
Results....258
Discussion....259
A.2 Adaptation....259
Analysis....260
Results....260
Discussion....261
A.3 General Discussion....262
Appendix B: Experimental stimuli....263
B.1 Stimuli of Experiment 1....263
Word targets....263
Pronounceable nonwords....264
Unpronounceable nonwords....264
B.2 Stimuli of Experiment 2....264

Word targets....264
Nonword targets....265
B.3 Stimuli of Experiment 3....266
Word targets....266
Nonword targets....267
B.4 Stimuli of Experiment 4....268
Word targets....268
Nonword targets....269
B.5 Stimuli of Experiment 5....270
Word targets....270
Nonword targets....271
B.6 Stimuli of Experiment 6....272
Word targets....272
Nonword targets....272
B.7 Stimuli of Experiment 7....273
Word targets....273
Nonword targets....274
B.8 Stimuli of Experiment 8....275
Word targets....275
Nonword targets....276
B.9 Stimuli of Experiment 9....277
Word targets....277
Nonword targets....278
B.10 Stimuli of Experiment 10....279
Word targets....279
Nonword targets....280
Appendix C: Generated items for assessing the OT3 metric....281
Appendix D: Default parameters of computational models....287
D.1 Bayesian Reader....287
D.2 Spatial Coding Model....287
Appendix E: Detailed results of the RT analyses in vincentiles....289
E.1 Analysis in deciles of Experiment 1....289
E.2 Analysis in deciles of Experiment 2....291
E.3 Analysis in deciles of Experiment 3....292
E.4 Analysis in deciles of Experiment 4....293
E.5 Analysis in deciles of Experiment 5....294
E.6 Analysis in deciles of Experiment 6....295
E.7 Analysis in deciles of Experiment 7....296
E.8 Analysis in deciles of Experiment 8....297
E.9 Analysis in deciles of Experiment 9....298
E.10 Analysis in deciles of Experiment 10....299

Index of Tables

Table 2.1....65
Table 2.2....66
Table 2.3....69
Table 2.4....73
Table 2.5....80
Table 2.6....82
Table 2.7....84
Table 2.8....85
Table 2.9....88
Table 3.1....97
Table 3.2....99
Table 4.1....108
Table 4.2....112
Table 4.3....114
Table 4.4....122
Table 4.5....126
Table 5.1....135
Table 5.2....137
Table 5.3....139
Table 5.4....142
Table 5.5....145
Table 5.6....146
Table 6.1....157
Table 6.2....158
Table 6.3....161
Table 6.4....162
Table 6.5....166
Table 6.6....167
Table 7.1....179
Table 7.2....180
Table 7.3....184
Table 7.4....186
Table 8.1....203
Table 8.2....204
Table 8.3....205
Table A.1....258
Table A.2....258
Table A.3....261
Table A.4....261
Table D.1....287
Table D.2....287
Table E.1....289
Table E.2....290
Table E.3....290
Table E.4....291
Table E.5....292

Table E.6.....292
Table E.7.....293
Table E.8.....293
Table E.9.....294
Table E.10.....294
Table E.11.....295
Table E.12.....295
Table E.13.....296
Table E.14.....296
Table E.15.....297
Table E.16.....297
Table E.17.....298
Table E.18.....298
Table E.19.....299
Table E.20.....299
Table E.21.....300

List of Figures

Figure 2.1....77
Figure 2.2....77
Figure 2.3....78
Figure 2.4....83
Figure 3.1....98
Figure 4.1....109
Figure 4.2....115
Figure 4.3....117
Figure 4.4....118
Figure 4.5....121
Figure 4.6....123
Figure 4.7....124
Figure 4.8....128
Figure 5.1....136
Figure 5.2....141
Figure 5.3....141
Figure 5.4....145
Figure 6.1....157
Figure 6.2....161
Figure 6.3....166
Figure 7.1....179
Figure 7.2....185
Figure 8.1....194
Figure 8.2....194
Figure 9.1....215

0. Preface

This thesis investigates the early stages of visual word recognition. In an interactive activation framework (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982) each word is represented by a node. All word nodes form the mental lexicon. In recognising a word, several word nodes are activated, but only one word node represents the stimulus most accurately. Thus, the interaction between these word nodes is crucial for establishing which word node provides the best match for a given stimulus, both in computational models (Davis, 1999, 2010), and in humans.

The experiments investigated how this selection is achieved and the way in which word nodes interact with each other. Masked priming studies (see section 1.1) have extensively investigated this interaction using related primes and targets, e.g. repetition priming (e.g., Forster & Davis, 1984), semantic priming (e.g., Perea & Gotor, 1997) and form priming (Davis & Lupker, 2006; Ferrand & Grainger, 1992; Forster, Davis, Schoknecht, & Carter, 1987). These are cases where the critical prime and target nodes form one cluster within the lexicon. In computational modelling, the experimental results were explained by homogeneous inhibition processes where activation in one word node can suppress activation in all other word nodes (e.g., Davis, 2010; Grainger & Jacobs, 1996; McClelland & Rumelhart, 1981). In reading normal text, words are read in fast succession and a strongly activated word node (from identifying the previous word) could interfere with recognising the next word (Grainger & Jacobs, 1999). One way of testing how two unrelated items interact, is investigating response congruency priming effects. Response congruency refers to a situation where a prime elicits an implicit response that is the same as the explicit response required to the target. A word target preceded by a word prime could result in a delayed response due to interference or in an advantage due to increased lexical activity. The impact of a prime on the lexical decision latency has been established with related primes, but the evidence for effects using unrelated primes is still inconclusive (Norris & Kinoshita, 2008; Perea, Fernández, & Rosa, 1998; Perea, Gómez, & Fraga, 2010). If response congruency effects can emerge, this would shed light on the mechanisms in the

lexicon, specifically on how activation in two unrelated word nodes is handled by the recognition system.

The experiments in this thesis were concerned with response congruency effects. Experimental data were also simulated in computational models to establish whether these models can account for the pattern of results.

In order to investigate the exact conditions where response congruency effects can emerge an effective way of manipulating the amount of lexical activity that the primes and the targets trigger was required. Thus, an algorithm for generating nonwords with a specific level orthographic typicality was developed and tested in a separate lexical decision experiment.

- Chapter 1 reviews the literature on response congruency effects and other potential sources of the effect that were not related to the lexicon and thus, had to be carefully avoided in the experiments.

- Chapter 2 introduces an algorithm for measuring orthographic typicality and empirical evidence testing the predictions of the algorithm's metric.

- Chapter 3 presents the first experiment and establishes that response congruency effects can emerge in lexical decision.

- Chapter 4 reviews the Bayesian Reader (Norris, 2006; Norris & Kinoshita, 2008) and the Spatial Coding Model (SCM, Davis, 2010) of word recognition. Also, a modification to the SCM is suggested and tested.

- Chapter 5 presents experiments that investigated how the task difficulty influences the presence or absence of the response congruency effect.

- In Chapter 6 the impact of the primes on response congruency was tested and it was established that the effect depends on the primes that were used in the experiment.

- Chapter 7 manipulates prime lexicality and investigated the source the response congruency effects.

- Chapter 8 presents simulations with the modified SCM testing the effectiveness of the suggested changes.

- Chapter 9 summarises the experimental and computational findings.

The findings of the experiments and the simulations showed that it is not the lexicality of the prime, but its orthographic typicality that can trigger response congruency effects. Furthermore, the evidence favoured a selective inhibition mechanism in the lexicon over homogeneous inhibition. That means only those word nodes that are related to each other compete for the best match to stimulus (selective inhibition), but not all word nodes (homogeneous inhibition).

1. Response congruency effects

Response congruency effects can occur in categorisation tasks. In such cases, a target stimulus receives a faster and more accurate response when it is preceded by an unrelated stimulus that elicits the same response compared to a stimulus that elicits a different response. This chapter explains the masked priming paradigm. Theoretical accounts to explain the underlying processes are reviewed and predictions specific to lexical decision experiments are derived. Finally, experiments where response congruency effects have been reported in lexical decision are reviewed.

1.1 Masked priming

The masked priming paradigm is one of the most popular experimental methodologies used today. In this experimental paradigm participants respond to a stimulus that is preceded by a briefly presented prime. Typically, the prime is masked by a preceding masking pattern and the target masks the prime backward (e.g., Forster & Davis, 1984). Also, the presentation duration is kept short (about 50 ms). Both a short presentation duration and the presence of preceding and succeeding stimuli contribute to masking the prime from conscious perception. Thus, participants cannot report the prime and are often unaware of its presence. The priming effect is measured by comparing a critical priming condition to a control condition within the same target stimulus (Forster, 1998).

A wealth of experimental findings has been produced using masked priming. Numerous concepts in cognition have been investigated, including semantics in numbers (Dehaene et al., 1998) and words (C. Brown & Hagoort, 1993; Greenwald, Draine, & Abrams, 1996), processing of orthographic form (Forster & Davis, 1991; Forster et al., 1987) and fearful behaviour (Siegel & Weinberger, 2009).

In particular in psycholinguistic research masked priming has been used extensively, for example in researching morphological processing (Frost, Forster, & Deutsch, 1997; Grainger, Colé, & Seguí, 1991; McCormick, Brysbaert, & Rastle, 2009), phonological effects in word recognition (Ferrand & Grainger, 1992, 1994; Lukatela, Frost, & Turvey,

1998) and in other research objectives (e.g., Kinoshita & Lupker, 2003). One very robust effect is the identity or repetition priming effect (e.g., Forster & Davis, 1984; Forster, Mohan, & Hector, 2003). The prime is a lowercase version of the target (*table* – *TABLE*) and is compared to unrelated primes (*drive* – *TABLE*). An advantage in reaction times (RT) and accuracy of the identity priming condition has been reported in lexical decision (Forster & Davis, 1984), in a recognition task (Evett & Humphreys, 1981; G. W. Humphreys, Quinlan, Evett, & Besner, 1987) and in semantic categorisation (Forster, 1985; Klauer, Eder, Greenwald, & Abrams, 2007). Priming effects have also been reported when the prime is associatively related to the target, e.g. in lexical decision, RTs are faster when *TABLE* is primed by *chair* than by an unrelated word like *drive* (Perea & Gotor, 1997; Sereno, 1991). Thus, masked primes influence the speed and accuracy of responses to target stimuli. An identity priming effect indicates that participants abstracted from the case difference between prime and target and processed abstract letter identities. The associative priming effect indicates that the prime was processed up to a relatively high level where semantic information was extracted.

1.1.1 Effective masking

Whether the primes in an experiment are perceived consciously or unconsciously depends on the strength of the prime stimulus. The strength of the prime depends on the presentation duration of the prime (e.g., G. W. Humphreys, Besner, & Quinlan, 1988; G. W. Humphreys et al., 1987) and whether it is masked (e.g., Dehaene et al., 2001). Dehaene et al. (2001) manipulated the masking condition by varying the order of a 70 ms masking pattern and a 70 ms blank following after the prime. Though the total duration of the trials and the prime duration was kept constant, the participants were able to detect, name and memorise the primes initially followed by the blank and then by the masking pattern. Participants were not able to detect, nor to report, nor to memorise the primes that were initially followed by the backward mask followed by a blank. The spatial location of prime and mask is also important. In the standard paradigm (Forster & Davis, 1984) the prime is covered by the forward and the backward mask. Though masking can be effective with adjacent stimuli in a metacontrast paradigm (Ansorge & Neumann, 2005; Kahneman, 1968). Under

conditions with a greater spatial offset the masking effect can break down and primes can be perceived consciously (e.g., Breitmeyer & Hanif, 2008). Varying the time span during which the prime is visible is also effective in manipulating conscious perception. The stimulus onset asynchrony (SOA) refers to the time span between prime and target onset. If the SOA is increased and no masking pattern is displayed between prime and target, the visibility of the prime increases. This can distinguish between conscious and unconscious perception and hence the effectiveness of masking provided by the following presentation of the target. Manipulating the SOA has resulted in qualitatively different behavioural data. For example, orthographic effects were reported with primes that were presented briefly and that were unconsciously perceived, but primes that were presented for longer and that were consciously perceived did not produce orthographic effects (G. W. Humphreys et al., 1987).

An effective masking is achieved by keeping the prime presentation duration short and by carefully placing the forward and backward mask.

1.1.2 Strategic effects

Hindering participants from identifying the prime consciously helps to reduce the effect of conscious systematic strategies. Forster (1998) found that participants showed a substantial expectancy effect in unmasked priming, i.e. consciously available primes, but this effect was reduced substantially in masked primes that were unavailable for conscious processing. On the other hand, experiments have shown that participants can exploit masked primes strategically to some extent. Bodner and Masson manipulated the proportion of primes with informative cues for the target compared to the unrelated primes in several studies. The results have shown that participants can adjust how they are processing the prime, even if they are not aware of their strategic prime evaluation in lexical decision (Bodner & Masson, 2001; Bodner, Masson, & Richard, 2006), semantic categorisation (Bodner & Masson, 2003) and naming (Bodner & Masson, 2004). In general a higher proportion of related primes increased the size of the priming effects. The masked priming procedure still allows to minimise strategic effects compared to unmasked priming. The effects described by Bodner and Masson

can be reduced by carefully choosing the priming conditions to balance informative and uninformative primes.

1.1.3 Masked versus unmasked priming

Priming effects differ between masked and unmasked priming. Segui and Grainger (1990) showed that word frequency interacts with SOA in lexical decision. Unmasked primes of a lower frequency than the target interfered with target identification, whereas masked primes of a higher frequency showed interference effects. Segui and Grainger (1990) attributed the effect in masked primes to early stages in processing and the effect in unmasked primes to later stages. Qualitative differences have also been reported in repetition priming (G. W. Humphreys et al., 1988), where repetition effects occurred in masked but not in unmasked primes, and orthographic priming (G. W. Humphreys et al., 1987), where orthographic effects were found in masked but not in unmasked primes. Unconscious and conscious processing of a stimulus has been investigated using neurophysiological data as well. C. Brown and Hagoort (1993) measured event-related potentials in a semantic categorisation task. They attributed a weaker N400 in masked primes compared to unmasked primes to a less extensive semantic integration in masked primes. An attenuated activation by masked compared to unmasked words has also been reported by Dehaene et al. (2001). Furthermore, Dehaene et al. (2001) showed a progressively decreasing pattern of activation along the path of word processing comparing masked to unmasked words. This suggests that the same mechanisms operate on masked and unmasked stimuli (Dehaene et al., 1998), but the processes are more or less incomplete with the masked stimuli. Finally, this difference in processing stages results in different effects depending on the priming procedure.

A major difference between masked and unmasked primes is the conscious perception of the unmasked stimuli. Humphreys et al. (1988; but also see Norris & Kinoshita, 2008) have suggested that masked primes are not perceived as two distinct objects by the recognition system. Thus, their processing is integrated as if it was one stimulus, which is not the case for unmasked primes that are perceived as two distinct objects. Another suggestion is that after an identification is completed the processing

units involved undergo a reset. This was the case in unmasked priming, but there was no reset in masked priming (Forster, 1999, 2009; Forster et al., 2003). In computational models (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Davis, 1999, 2010; Grainger & Jacobs, 1996; McClelland & Rumelhart, 1981) a similar account is accepted, when masked priming is achieved by preserving activation and unmasked priming is equivalent to processing two stimuli.

These findings suggest that the masked priming paradigm is effective in highlighting early stages of the recognition process whilst avoiding that participants make use of the prime in a conscious way. Unmasked priming is more likely to investigate later stages of processing and effects of conscious processing.

1.1.4 Prime duration

The prime presentation duration does not only affect the stimulus perceptibility, but also the amount of processing that the recognition system is allowed to perform before target onset. Varying the SOA and keeping the primes unconscious, phonological priming effects were reported using a 60 ms SOA, while a shorter SOA (about 30 ms) did not reveal phonological priming (Ferrand & Grainger, 1992, 1994). Similarly, in an ERP study the onset of phonological effects was measured later than the onset of orthographic effects (Grainger, Kiyonaga, & Holcomb, 2006). The emergence of phonological and orthographic priming effects is task dependent (Grainger & Ferrand, 1996). The size of the priming effect is also linked to the prime duration, e.g. Forster et al. (2003) have reported no difference between identity and form priming using neighbours with a prime duration of 20 ms and 30 ms. Using 50 ms prime duration the identity priming effect increased to about 50 ms, whereas the form priming effect remained static at about 30 ms. They suggested the prime duration forms a limit to the size of the priming effect. Though there is evidence that priming effects can be larger than the SOA with mixed case targets (Bodner & Masson, 1997), where repetition priming effects of 71 ms (high frequency word targets) to 93 ms (nonword targets) have been found with an SOA of 60 ms in a lexical decision task. The physical overlap between the prime and the target was investigated and said not to contribute to the results (Bodner & Masson, 1997, Experiment 2b). Though some doubts on the

effectiveness of the mask may remain. In general, the priming effect is bound to the magnitude of the prime duration.

Another aspect of timing is the persistence of the priming effect across time. The effect of masked primes lasts for a shorter duration than the effect of consciously perceived unmasked primes. Unmasked primes can leave a memory trace and influence the next prime-target pair. Masked primes appear to be limited in their effect if the SOA is greater than 150 ms (Greenwald et al., 1996). Though, the absence of positive priming might be due to a negative priming effect building up at SOAs around 100 ms (Eimer & Schlaghecken, 2003; Schlaghecken, Bowman, & Eimer, 2006) and it cannot be concluded that any effect of priming is absent after 100 ms. More recent findings suggest that the priming effects measured at different SOAs differ in their loci and with that in their prevalence at certain points in time, e.g., Forster (2009) has suggested that there are different sources for form and identity priming. He has reported form priming diminished with a masked letter string intervening between prime and target (resulting in about 100 ms SOA). The identity priming effect was reduced but still significant. When the intervening stimulus was unmasked by increasing the presentation duration to 500 ms (about 550 ms SOA) form priming was unaffected compared to trials without intervening stimuli. But with the unmasked intervening stimulus the identity priming effect was reduced to the level of form priming. Forster (2009) concluded that there are two distinct sources of the priming effects and suggested the effect with unmasked intervening stimuli is located in a semantic level and the effect with masked intervening stimuli in a form level. The shorter SOA (no form priming, but identity priming) is then related to early processing stage on form level and the longer SOA with semantics at a later processing stage. This finding is pointing in a similar direction as the results by Ferrand and Grainger (1992, 1994), where shorter SOAs were associated with earlier processing stages.

1.1.5 Conclusion

Priming is effective in shedding light on early processing stages of word recognition. In order to hinder conscious perception of the primes the SOA has to be short. Also, the forward and backward masks themselves have be effective either by ensuring the

prime is fully covered by the masks or by using an effective metacontrast mask. Preventing the primes from conscious perception is useful for hindering participants from becoming aware of the experimental conditions and from employing a strategy. Even with appropriate masking, effects of unconsciously applied strategies can emerge, but balancing the proportion of related and unrelated primes can avoid them.

1.2 Theoretical accounts of response congruency effects

Response congruency effects could occur in any categorisation task and effects have been reported in numerous studies. For example in a number magnitude estimation task the numbers 1 to 4 are categorised 'smaller than five' and 6 to 9 'larger than five'. Participants are faster to categorise *4* as 'small' when primed with a small numeral or digit such as *two* compared to a large prime such as *seven*, i.e. the sequence *two – 4* is associated with faster responses than *seven – 4* (e.g., Dehaene et al., 1998; Kinoshita & Hunt, 2008; Kunde, Kiesel, & Hoffmann, 2003; Naccache & Dehaene, 2001). Similar congruency effects were reported with a variety of different types of stimuli, including narrow categories semantic categorisation (Forster, 2004, Experiment 4), valence categorisation (Chan, Ybarra, & Schwarz, 2006) and gender categorisation (Klauer et al., 2007, Experiment 3).

The evidence shows that response congruency effects can emerge in tasks that involve printed word stimuli including semantic categorisation tasks. I continue reviewing theoretical accounts.

1.2.1 Depth of processing

1.2.1.1 Deep Processing

In a deep processing account (Dehaene et al., 1998) the prime is analysed with regards to task instructions and processed up to the motor level. In a review, Kunde et al. (2003) coined the term 'elaborate processing'. The response congruency effect is attributed to interference on the motor level. Dehaene et al. (1998) did two number magnitude estimation experiments and collected behavioural, ERP and fMRI data. All measures confirmed that there was a response congruency effect. In incongruent trials

where the response towards the prime is opposite the target response, a short-lived lateralised readiness potential was found. This was said to reflect a prime-related response and was followed by a stronger target related signal on the opposite side of the cortex. In the congruent condition this change in the signal was absent. The data from the fMRI study located the areas where the congruency effect emerged in the motor cortex. A stimuli inherent effect like associating numbers with a spatial orientation (Brysbaert, 1995) can be excluded, because the response hands were varied systematically. From the behavioural, electrical and haemodynamic data Dehaene et al. (1998) concluded that the primes are processed unconsciously up to the point of response initiation. In their account the prime is perceived and participants apply task instructions unconsciously to the prime. The processing of the prime is 'deep' going through semantic evaluation with regards to the instructions and to the respective response. Eventually, a readiness potential in the respective motor cortex is triggered. If the readiness potential triggered by the prime response is incompatible with the actual target related response this will cause a conflict. Resolving this conflict takes time and thus, incongruent trials will be responded to slower than trials in which there is no response conflict. Congruent trials also benefit from the fact that the correct response has already received motor activation due to the prime. The combination of inhibition and facilitation means that congruent responses are faster than incongruent responses. All these processes are assumed to be unconscious to the participant.

1.2.1.2 Stimulus-response mapping

Instead of explaining the data by assuming a deep processing of the prime, Damian (2001) explained the response congruency effect as a result of repetition learning. Damian (2001) asked participants to judge whether the object denoted by the presented word was larger or smaller than a reference square presented on screen. In the first experiment a limited stimulus set was used and repeated frequently as primes and targets. The response priming effect increased with the number of blocks. Damian (2001) argued the increase reflected a stimulus-response mapping building up as the experiment progressed. In Experiment 2 the primes were replaced by a set of nouns that did not occur in the response set. Under these conditions the response

congruency effect was absent. Damian (2001) concluded that responding to the same stimulus repeatedly results in a stimulus-response mapping. Once this mapping is established the participants do not need to process the meaning of the stimulus, but can instead make use of their previous judgement. The data showed that stimulus-response mappings may develop after only a very few trials. This finding challenges the elaborate processing account proposed by Dehaene et al. (1998). Activity in the motor cortex shortly after prime presentation could be the result of activating a mapping formed earlier in the experiment rather than the result of semantic processing.

1.2.1.3 Prime novelty

In a response to the stimulus-response mapping account, that questioned Dehaene et al.'s (1998) finding, Naccache and Dehaene (2001) performed an experiment using novel primes. According to the stimulus-response mapping account only primes that the participants encountered before should be effective, whereas novel primes should not. Naccache and Dehaene (2001) conducted a number magnitude estimation experiment varying prime novelty and prime notation as Arabic number and numeral systematically. A congruency effect was observed independently of the prime novelty and the notation format. The independence from prime novelty was interpreted as support for semantic processing and the elaborate processing account (Dehaene et al., 1998), because a simple mapping from a specific stimulus to a response cannot explain the data (Naccache & Dehaene, 2001). This point is strengthened by a number of further studies reporting effects with novel primes (e.g., Greenwald, Abrams, Naccache, & Dehaene, 2003; Kinoshita & Hunt, 2008; Klauer et al., 2007; Norris & Kinoshita, 2008; Quinn & Kinoshita, 2008; Reynvoet, Gevers, & Caessens, 2005).

1.2.1.4 Summary

Response congruency effects are found regardless of whether participants have made conscious responses to the prime stimuli. The absence of congruency priming effects in Damian's (2001) experiment could be due to task differences and I will discuss this in section 1.3.

1.2.2 Processes influencing priming effects

1.2.2.1 Action triggers

Kunde et al. (2003) argued that the effectiveness of the primes depends on the content of 'action triggers'. The participants form them in accordance with the task instructions and adjust them while performing the task in training and throughout the experiment. In other words in number magnitude estimation tasks participants are prepared to respond to the items accordingly – even if they have not been presented with any stimuli. If the instructions do not mention this particular semantic dimension, the action triggers will not contain this information. Under these circumstances a congruency priming effect is not expected in an action trigger account. But the deep processing account that assumes automatic semantic processing predicts a congruency priming effect.

Kunde et al. (2003) instructed participants to judge the number magnitude with respect to *5*. Only a subset of all digits (*2* and *4* → *smaller* response, *7* and *9* → *greater* response) was presented to participants, but a response congruency effect with old (*2, 4, 7* and *9*) and novel primes (*1, 3, 6* and *8*) was present (Kunde et al., 2003, Experiment 1). However, the effect was absent when the same key assignment was used (*2* and *4 left* response, *7* and *9 right* response), but the instructions strictly avoided mentioning the semantic context of magnitude (Kunde et al., 2003, Experiment 3). In contrast to the semantic task, the simple mapping task revealed facilitatory priming for old primes (*2, 4, 7* and *9*) and an inhibitory priming effect in novel primes (*1, 3, 6* and *8*). Kunde et al. (2003) argued that the absence of priming in novel primes with the non-semantic instructions is a challenge for the deep semantic processing account. In such an account the semantics of the prime are analysed automatically and a congruency priming effect should occur regardless of task instructions. On the other hand Dehaene et al. (1998) argued the task instructions would be applied to the primes, which is not excluded by the results. Due to the way the numbers were assigned (either *2, 4, 7* and *9* or *1, 3, 6* and *8* in the old set), some participants reported having interpreted the task as an even-odd distinction (Kunde et al., 2003). If so, the results suggest the task instructions were applied to the primes accordingly. In old primes a facilitatory effect

with regards to the magnitude reference *5* was measured, but this is indistinguishable from a positive effect in an even-odd distinction task as the two properties always coincided. The negative priming effect with regards to the magnitude task in novel primes is actually a positive effect in an even-odd distinction task. The positive effect is expected and the even-odd reinterpretation provides a good explanation for the findings – both the deep processing and the action trigger account are in line with this finding. Hence, the deep processing account cannot be discarded, if participants who were completely unaware of the actual semantic dimension just performed some other reasonable task.

The presence of a congruency effect in novel primes with or without semantic instructions indicates that stimulus-response mapping is not the only source of response congruency effects, in accord with Naccache and Dehaene (2001)'s finding. On the other hand, this empirical evidence cannot distinguish between the action trigger and deep processing account.

The predictions of the deep processing account and the action trigger account appear to be hard to distinguish. But, according to the action trigger account the participants prepare themselves to respond to the upcoming items, which is achievable in small sets e.g., in number magnitude estimation tasks the item set is usually small, with eight critical stimuli and *5* as the neutral item. In larger categories where items are not predictable the formation of action triggers is unlikely (see Klauer et al., 2007), but a semantic analysis is not excluded. In a size estimation task the stimuli set is large and hard to predict and a response congruency effect was not reported by Damian (2001). This suggests that participants could not form action triggers and with respect to the deep processing account that the semantic processing was too complex for a congruency effect to emerge.

1.2.2.2 Two processes model

If response congruency priming arises as a result of semantic activation, as in the deep processing account, this effect should occur in processing words other than numerals. If priming is the result of forming action triggers, then an effect is not

predicted in a large category. Klauer et al. (2007) conducted a valence classification experiment with novel primes where participants had to indicate whether the target is positive or negative and a gender classification experiment where participants had to indicate whether a first name target refers to a female or male person. Both tasks used novel primes and presumably a large category. In both tasks a priming effect was found. Klauer et al. (2007) concluded that at least two processes are involved in priming. First, a central priming component that is assumed to be semantic in nature. Secondly, a response related priming component that is similar to stimulus-response mapping (Damian, 2001) and action triggers (Kunde et al., 2003).

1.2.2.3 Conclusion

Response congruency effects can emerge from various sources. One source of priming effects is stimulus-response mapping (Damian, 2001) or as Klauer et al. (2007) phrased it response-related priming. Effects on the basis of this response component are likely to occur in motor cortex as a conflict between response options (Dehaene et al., 1998). On the other hand, these effects are not informative about an evaluative decision process. In restricting research to novel primes there are still at least two components that could result in a congruency priming effect. First, the participant can prepare themselves for the task and form action trigger sets (Kunde et al., 2003). The information in the action triggers maps stimuli to a response, so that interference will occur at response level. Secondly, a truly evaluative process can be biased by the influence of a prime. In order to isolate the latter effect the primes must be novel and the categorisation task sufficiently complex to avoid the formation of action triggers.

1.2.3 Summary

Response congruency priming effects with novel primes have been reported in number magnitude estimation (Naccache & Dehaene, 2001) and valence classification (Klauer et al., 2007) as well as in other tasks. On the other hand these effects have not been reported in size estimation (Damian, 2001) or as will be seen later in lexical decision with novel primes (Norris & Kinoshita, 2008; Perea et al., 1998, 2010). One dimension where these tasks might differ systematically is category size or semantic

coherence within categories. I review research with regards to category size and other effects of the experimental design in the following section.

1.3 Properties of the stimuli and the experiment

1.3.1 Category size

1.3.1.1 Cardinality

The effect of category size has been investigated by using masked primed semantic categorisation tasks. Analysing the *no*-responses to nonexemplar stimuli in a categorisation task Forster (2004, Experiment 1-2) found that in small categories (e.g., body parts) an effect of lexical frequency of the target stimulus is absent, but it is present in large categories (e.g. animals, Forster, 2004, Experiment 3). Also, nonexemplar targets can be primed by nonexemplar primes in large categories (Forster et al., 2003), whereas in small categories a response priming occurs in both, exemplar and nonexemplar targets (Forster, 2004, Experiment 4). Forster (2004) concluded that two different mechanisms operate in parallel and depending on category size the consequences of the one or the other become apparent. In small categories participants can predefine a set which is rapidly scanned for the target. This idea is related to the action trigger account (Kunde et al., 2003), where in small target sets responses are anticipated. Both, the serial search and the action trigger account produce similar predictions in small categories. The quick scan of a small response set allows to bypass lexical access and a frequency effect is absent. In contrast, in a large category such a list cannot be prepared and semantic properties are retrieved by means of lexical access. Thus, a frequency effect is found in large categories.

In a large category, exemplars and nonexemplars are accessed through the lexicon. In order to explain the small priming effect in nonexemplars in large categories Forster et al. (2003) suggested that accessing the lexical representation of the exemplar prime in an incongruent trial needs a resolution and in turn slows down incongruent compared to congruent responses. Though this mechanism would predict priming in exemplars as well, which is not consistent with the data and thus, the data provide a challenge for the entry opening model (Forster et al., 2003).

1.3.1.2 Category coherence

Quinn and Kinoshita (2008) investigated the effects of category size using a different definition of size. In their account, the category size is not determined by the number of items, but by the consistency in the shared features of its members. I will use Quinn and Kinoshita's (2008) terminology and refer to a category that is formed by a core set of semantic features shared by most of its members as *narrow* and to a category that is more loosely defined as *broad*. Specifically in broad categories, a case where one member of the set shares no feature with another member can occur. As a result, membership in a category cannot be defined by the means of shared features, but is rather defined by family similarity. Wittgenstein (1989/1914) pointed this out in his philosophical account using the example of *game* and showed that there is presumably no feature shared by all entities that can be referred to as a *game*. In Forster's (2004) category search account the number of members is crucial because it determines whether an exhaustive search through its members or lexical access is performed. In contrast, the breadth of a category does not necessarily limit the number of its members.

In a semantic categorisation experiment Quinn and Kinoshita (2008) used the broad category of animals and manipulated the number of shared features and category congruency between primes and targets. The results showed that the semantic primes resulted in priming effects compared to category congruent primes in both exemplar (animal: *hawk/mole – EAGLE*) and nonexemplar targets (animal: *pistol/boots – RIFLE*). Response congruency showed a strong tendency in nonexemplars (animal: *boots/camel – RIFLE*), but no effect in exemplars (animal: *mole/knee – EAGLE*). The results from the experiment using narrow categories showed a response congruency effect in exemplar targets (body parts: *arms/taxi – HAND*). In nonexemplar targets a congruency effect occurred that was independent of semantic relatedness of prime and target (planets: *pistol=boots/mars – RIFLE*). Furthermore, in nonexemplar targets a frequency effect occurred in both narrow and broad categories. This frequency effect seems to rule out a category search account where no frequency effect was predicted in nonexemplars in small categories, which was in line with prior data (Forster, 2004).

Quinn and Kinoshita (2008) suggest that the primes activated semantic features and that participants used different strategies to monitor them. In narrow categories, the relevant set of features is small and irrelevant features are ignored. Restricting the set of monitored features in narrow categories is related to the idea of forming action triggers (Kunde et al., 2003). Whenever one of the predetermined features is activated the respective response is triggered. In a broad category, the set of critical features is not restricted as explained above (Wittgenstein, 1989/1914) and as a result all features are monitored. This explains why in an animal categorisation task semantically related primes elicit stronger congruency priming effects than category congruent primes (animal: *pistol>boots – RIFLE*) and the absence of the effect in narrow categories (month: *pistol=boots – RIFLE*).

The semantic feature overlap hypothesis assumes that various features are monitored in a broad category semantic categorisation task, whereas a more restricted set is monitored in narrow categories. Crucially in this account, these features have to be connected to the representation of the word naturally driven by word meaning rather than by arbitrary task requirements in the experiment. The authors assume that there is a natural set of features that constitutes body parts or animals. Relying on these connections a priming effect can emerge. In contrast, Quinn and Kinoshita (2008) argue that no priming effect emerges where no natural sets of features exist defining the category, e.g. in ad hoc categories. Examples of ad hoc categories include *things that could fall on your head* (Barsalou, 1983), *possible gifts* (Barsalou, 1982) and *things that are larger than a 20 by 20 cm square reference* (Damian, 2001). The point relates to Forster's (2004) restriction that priming is only predicted in an automatically processed task, excluding size estimation. Furthermore, this restriction can also be thought of as a restriction on the applicability of action triggers. As mentioned before action triggers are unlikely with large categories (Klauer et al., 2007), but this may refer to category coherence, e.g. numbers in a magnitude estimation task with reference to *5000* are as coherent as with reference to *5* despite set sizes differing dramatically.

1.3.1.3 Impostors

Impostors are words that share central features with a category, but are not members of this category (e.g., *mind* with respect to *body parts*). If an overlap in semantic features is necessary to produce a response congruency effect, but not the category membership per se, then impostors should be effective primes in a categorisation task. Quinn and Kinoshita's (2008, Experiment 4) data confirm that impostor primes show effects that are more similar to an exemplar prime than to a nonexemplar prime. In nonexemplar targets the impostor primes did not differ from exemplar primes (body part: *mind=ear – LAKE*) in showing a congruency effect compared to unrelated nonexemplar primes (body part: *door – LAKE*). In exemplar targets the impostor primes (body part: *mind – HEAD*) facilitated the response compared to unrelated nonexemplar primes (body part: *door – HEAD*), but the facilitation was less strong compared to exemplar primes (body part: *ear – HEAD*). This is in line with the prediction that semantic features are the basis for congruency effects in semantic categorisation. Also, participants showed high error rates in classifying impostors in an unprimed speeded categorisation task (30.4%). The error rate dropped dramatically in an unspeeded task (1.8%) showing that participants were able to classify the items correctly.

The impostor effect can also be transferred to a lexical decision task. Nonwords can vary in the degree in which they resemble words in a particular language (Coltheart, Davelaar, Jonasson, & Besner, 1977; Duyck, Desmet, Verbeke, & Brysbaert, 2004; Hauk et al., 2006; Westbury & Buchanan, 2002). Nonwords in a lexical decision task resemble the nonexemplar targets in a semantic categorisation task as the critical features are not activated. Words resemble the exemplar targets with activation in semantic features. An impostor in a lexical decision task can be a nonword that is sufficiently wordlike to facilitate a word-response rather than a nonword-response. Showing such an impostor effect in a lexical decision task would show that those features that differ between typical nonwords and impostor nonwords are very important in the early stages of lexical access. I will return to this point in Chapter 7.

1.3.2 Task expertise

Kiesel, Kunde, Pohl, Berner, and Hoffmann (2009) have reported a masked priming study where participants had to judge whether a three by three fields chess situation was depicting a checking or not. The king was positioned in the upper left corner in all displays and the attackers (either rook or knight) varied among four positions in total. Only those participants who had expertise in chess playing showed a response congruency effect, whereas those who had just started playing chess did not. This finding might imply that a semantic representation has to evolve over time or that automaticity evolves with practice (Damian, 2001). Also, the absence of response congruency priming in tasks where automaticity has not evolved supports the assumption that automaticity is a prerequisite to find response congruency effects (Forster, 2004). This contrasts with e.g., linguistic stimuli where the extraction of meaning appears to be automatic and can hardly be suppressed (e.g., Stroop, 1992/1935). Another interpretation is that novice participants cannot prepare themselves for the task and therefore fail to form appropriate action triggers (Kunde et al., 2003) even though the number and the sort of displays is limited. Also, a response related process will not affect decision speed or accuracy in novel primes (Klauer et al., 2007).

Independent of the formulation, the participants need a certain level of familiarity with the categorisation task for a response congruency effect to emerge. A size estimation task (Damian, 2001) might involve this problem, where participants may never have thought about the size of an object in relation to an arbitrary reference.

1.3.3 Response speed

Prime-induced congruency effects have been reported to be strongest in fast RTs (Abrams, 2005; Burle, Possamaï, Vidal, Bonnet, & Hasbroucq, 2002; Greenwald et al., 2003; Kinoshita & Hunt, 2008). Thus, the differences reported with expert and novice participants (Kiesel et al., 2009) can also be interpreted as a matter of the speed with which the task was completed. In this respect, it is worth noting that participants with

no training, but fast RTs showed response congruency priming effects in other tasks (Forster et al., 2003; Klauer et al., 2007; Kunde et al., 2003; Perea & Gotor, 1997).

The RT in a task is linked to the difficulty of the task. In a hard task slow RTs are expected and with this no response congruency effect. This is also related to action triggers and automaticity, where fast responses and response congruency effects are expected. Three different kinds of difficulty can arise. First, the two categories are just hard to distinguish because members are very similar. In such cases the responses will be slow, even if only part of the list is hard to categorise (e.g., Dorfman & Glanzer, 1988). This reflects a difficulty due to the selected items in the list and it is independent of participants' prior experience or the automaticity in the task. Secondly, a task might refer to a very few, clearly distinguishable semantic features, but some participants might not be able to extract this information from the stimuli quickly, e.g. in the chess experiment (Kiesel et al., 2009). In this case only participants with experience showed a response congruency effect. Turning this around, the participants' experience in a categorisation can be assessed by the emergence of a response congruency effect. Finally, if the categories are very broad and the criteria not strictly assessable, the task can be perceived as hard. Especially, if there are a number of uncertain cases, e.g. in *things to take from one's home during a fire* (Barsalou, 1983). In these cases a priming effect is very unlikely. One reason might be an unsystematic disagreement of the correct categorisation between experimenter, participants and within participants. Another, that these tasks will require quite some training on a somewhat uncertain criterion.

In summary, there are three main sources of task difficulty: very similar categories independent of the individual participants, difficulty of extracting the required information depending on individual experience, and difficulty of classification depending on the task as such.

For the experiments in the following chapters ensuring that there is agreement about the categorisation of the stimuli is important. This was achieved by checking the frequency of word targets and the number of correct responses these stimuli received

in a mega-study (Balota et al., 2007). Finally, the similarity between the stimuli in each category can be manipulated. I will make use of this technique in order to manipulate the task difficulty between experiments.

1.3.4 Conclusion

In small categories, participants can concentrate on a few distinctive features and bypass lexical access. Theoretical accounts such as the action trigger, entry opening and semantic overlap account can explain the respective empirical data. In large categories the entry opening account predicts congruency priming in nonexemplar targets in large categories as a result of interference at form level. The action trigger account cannot be applied to large categories. Also, a response congruency effect is not predicted by the semantic overlap account, because all features are monitored and overlapping features are effective in exemplar and nonexemplar targets. This account is also supported by the data from impostor primes, because the activation of some features can drive congruency effects. Furthermore, response congruency effects can only arise if the participants have some experience with the task and response congruency effects are more likely in fast responses.

1.4 Response congruency effects in lexical decision

In a masked primed lexical decision task the participants are asked to indicate whether the target is a word or a nonword (Forster & Davis, 1984). The lexical decision task has been criticised and identified as artificial (Forster et al., 2003), but the task is performed in every day life. Chaffin, R. K. Morris, and Seely (2001) pointed out that normal readers encounter numerous words, acronyms or slang expressions that they have not seen before. In order to associate them with a new meaning and prevent a misidentification they had to be recognised as new. For school children in the final years the estimate is about 10 to 15 new words per day (Landauer & Dumais, 1997). Perhaps, in adult life the estimate is higher for periods of changing life circumstances and lower in other periods. Importantly, the identification as new is similar to a lexical decision task. Thus, it could be argued that there is a natural mechanism to work out whether a letter string is known or unknown which is exactly what a lexical decision requires.

In the masked priming paradigm the difference between two or more priming conditions is of interest and thus, the influence of the prime on the explicit task. The participants are not responding to the prime, but this implicit response is of specific interest. If prime and target are of the same lexicality the responses are congruent (*word – WORD, nonword – NONWORD;* e.g. *crown – QUIET, apvxa – MIYTD*) and otherwise they are incongruent (*nonword – WORD, word – NONWORD;* e.g. *apvxa – QUIET, crown – MIYTD*). A difference between congruent and incongruent trials would constitute a response congruency effect.

As noted earlier, participants can extract semantic information from primes, giving rise to response congruency effects in semantic categorisation (e.g., Forster et al., 2003). Semantic effects have been reported in lexical decision as well (Perea & Gotor, 1997; Sereno, 1991) indicating that the processing of the prime's meaning influenced the decision process. Determining the lexicality of a stimulus requires less specific information that could be available earlier in processing. Thus, it is possible that information about prime lexicality is available, which could contribute to a response congruency effect in lexical decision.

1.4.1 Predictions from theoretical accounts

Theoretical accounts of congruency priming have different predictions on lexical decision. I review the accounts introduced in this chapter and derive predictions for lexical decision experiments.

1.4.1.1 Deep processing and stimulus-response mapping

The deep processing account (Dehaene et al., 1998) predicts that participants apply the instructions to the prime and hence, a response congruency effect can emerge in a lexical decision task. This contrasts with the stimulus-response mapping account (Damian, 2001). A response congruency effect in a lexical decision task is very unlikely, because the stimuli have to become associated to their respective response. Learning the mappings seems to be more effective with increasing number of repetitions, but may also be fostered by small sets of stimuli. In a lexical decision experiment, the number of items per category is typically greater than 50 compared to 12 per category

in Damian's (2001) experiments. Also, the stimuli are typically not used repeatedly as targets. Thus, the predictions of the deep processing and the stimulus-response mapping account can be distinguished in novel primes where a deep processing account predicts a congruency effect but stimulus-response mapping does not.

1.4.1.2 Category size

The category size or coherence is important for the serial search (Forster, 2004), the semantic overlap (Quinn & Kinoshita, 2008) and action trigger account (Kunde et al., 2003), because it determines which mechanism is applicable and the predicted outcome.

In small categories, participants can form action triggers and a response congruency priming effect would be attributed to these prepared mapping relations (Kunde et al., 2003). In a lexical decision task, the categories are large. The number of words is a few ten-thousand compared to the number of months or the number of planets in our solar system. The category of words is also semantically broad, because every possible word meaning is part of the category. Thus, the formation of action triggers is unlikely and this account would not predict a response congruency effect in lexical decision.

The category search account (Forster, 2004) predicts that there is a congruency priming effect in *no*-responses in categorisations with large categories. Though words being a large category, Forster (2004) argues that lexical decision is not performed automatically which was a necessary condition for priming effects to emerge. The lack of automaticity is said to explain the absence of response congruency effects in lexical decision (e.g., Perea et al., 1998), but also in ad-hoc categories (e.g., Damian, 2001).

Although the semantic feature overlap account (Quinn & Kinoshita, 2008) refers to category size as well, the category search and the semantic feature overlap account differ with respect to lexical decision. Quinn and Kinoshita (2008) argue that congruency priming only occurs in natural categories. In broad categories, various features are monitored. By assuming all features can be monitored, the categories of words and nonwords can be distinguished by their total semantic activation. With this

assumption the category of words would form the broadest possible category. In turn, a response congruency effect was predicted in lexical decision as a result of interference and facilitation in semantic features. Showing a response congruency priming effect in lexical decision would provide evidence for the broad category of words in a semantic feature overlap account and challenge the prediction by the category search model.

1.4.2 Empirical evidence

1.4.2.1 Studies reporting null effects

Prior studies specifically aimed at finding a response congruency effect in lexical decision have not reported significant results in Spanish (Perea et al., 1998, 2010) or in English (Norris & Kinoshita, 2008). An inspection of the items used in Norris & Kinoshita (2008) showed that the nonwords were derived from words in the experimental set by exchanging and transposing letters. Also, the nonwords were all legal and formed wordlike stimuli. Thus, it could be that the words and nonwords in the experiment were very similar and the difficulty of distinguishing between very similar stimuli hindered a response congruency effect to emerge. The experimental data revealed a frequency effect in word targets, but there was no sign of a response congruency effect in neither word or nonword targets. Another study that is relevant, but not specifically aimed to find a response congruency effect was conducted by Sereno (1991). Two masked primed lexical decision experiments (Experiment 1 and 3) were designed to test graphemic and associative priming. There were two control conditions using unrelated word and unrelated nonword primes. The results showed that the conditions did not significantly differ, i.e. there was no congruency effect in word or nonword targets. In summary, the absence of a response congruency effect in lexical decision has been reported by several studies.

1.4.2.2 Studies reporting an effect

Klinger, Burton and Pitts (2000, Experiment 2) have reported a response congruency effect in a masked primed lexical decision task, both in word and nonword targets. An important aspect of this experiment is the practice procedure. In order to train

participants to respond within a response window they were exposed to all targets four times prior to the experimental block. Thus, participants may have formed stimulus-response mappings for the target stimuli during the practice blocks as suggested by Damian (2001). Since the primes were drawn from the same pool as the targets, the response congruency effect may have emerged as a result of the mapping relation. The conclusion that the effect is due to stimulus-response mappings is supported by the results of a semantic categorisation experiment where the same training procedure was employed (Klinger et al., 2000, Experiment 1). Next to the response congruency effect no other priming effects were reported. This is in contrast to other data where related nonexemplar primes facilitated responses to nonexemplar targets more effectively than unrelated nonexemplar primes (e.g., Quinn & Kinoshita, 2008). Thus, it is possible that in the semantic categorisation task participants did not process the stimuli semantically but relied on their previously learnt mappings. Similarly, stimulus-response mappings may have triggered the response congruency effect in the lexical decision experiment (Klinger et al., 2000, Experiment 2) because the same training procedure was applied. In contrast, Norris and Kinoshita (2008) who did not observe a congruency effect used all novel primes and the absence of a congruency effect in their data could be attributed to this difference in the experimental procedure.

Another study pointing to a possible congruency effect was conducted by Jacobs, Grainger, & Ferrand (1995, Experiment 2), where prime novelty was added as an additional factor in the analysis. In the experiment, the prime visibility was increased in four steps. The baseline was defined at a visibility level where none of the conditions elicited different response times, including an all-letters-different and an identity priming condition. In word targets, the two highest prime visibility conditions showed a significant difference between the unrelated word and unrelated nonword priming conditions. In nonword targets, the difference between unrelated word and unrelated nonword primes was significant in only one visibility condition. This finding was interpreted as a motor priming effect stemming from the activation in congruent and incongruent response codes. Jacobs et al. (1995) concluded that there was no evidence for the contribution of activation in lexical representations, because the response congruency effect occurred in both word and nonword targets. The analysis of prime

novelty did not reveal any difference in the response congruency effect, though this would have supported the hypothesis that the effect was due to interference at motor level. Furthermore, the absence of a prime novelty effect suggests that priming in lexical decision is not affected by very few conscious responses to the prime as a target. The absence of novelty effects could also reflect a smaller number of repetitions than in Damian (2001). Thus, in contrast to Klinger et al.'s (2000) results the findings of Jacobs et al. (1995) cannot be explained in terms of stimulus-response mappings. Otherwise a stronger congruency effect should emerged in old compared to novel primes. The data suggest that the primes influenced the processing of the target, e.g. by biasing the response to the target or by influencing decision channels.

Finally, the data collected by C. J. Davis and Lupker (2006, Experiment 1) indicated a response congruency effect, although C. J. Davis and Lupker (2006) did not discuss response congruency effects, their analysis indicated that a response congruency effect was observed in low and high frequency word as well as in nonword targets.

1.4.2.3 Empirical discrepancy

In summary, the empirical evidence is inconclusive. There are several studies that did not report a response congruency effect in lexical decision. Three studies showed an effect, but one could reflect the effect of prime stimuli that were repeatedly presented as targets.

Comparing the masked primed lexical decision studies that reported a response congruency effect and those that did not, one potential difference could be the mean RTs. As noted earlier, response congruency effects are more likely in faster than in slower RTs (e.g., Burle et al., 2002). Jacobs et al. (1995) who observed an effect of about 20 ms reported mean RTs ranging from 399 ms to 515 ms (Experiment 2). C. J. Davis and Lupker (2006) reported mean RT ranging from 571 ms to 757 ms and an effect of about 10 ms, supporting the assumption that the effect size might decrease in slow responses. Compared to Jacobs et al.'s (1995) data, the RTs of experiments not revealing an effect were considerably slower. Norris and Kinoshita (2008) report mean RTs between 526 ms and 682 ms (Experiment 1). The studies in Spanish reported 676

ms to 811 ms (Perea et al., 1998) and 583 ms to 695 ms (Perea et al., 2010). That means the slowest RT in Jacobs et al.'s (1995) experiment where an effect was reported is still faster than the fastest mean RT in experiments showing null effects. Nevertheless, these RT are comparable to C. J. Davis and Lupker's (2006) data and this indicates that the response speed might be a factor amongst others contributing to the presence of response congruency effects. The impact of the response speed could indicate that response congruency effects are short-lived and the leaky accumulator model (Usher & McClelland, 2001) could accommodate this. In this account the prime adds evidence in one of the decision channels, but since the accumulator is leaky this evidence trickles out as time continues. A larger response congruency priming effect would emerge in fast responses where the impact of the prime is more present than in slower trials. There are two critical factors. First, the difficulty of the task, because a hard task slows down responses and increases the effect of the leakage and secondly, the amount of evidence fed into the accumulators is critical. The greater the evidence added by the prime, the less vulnerable is the response congruency to the leakage. In summary, a leaky accumulator model could capture the empirical data introduced so far.

1.4.3 Conclusions

The empirical evidence in masked primed lexical decision is unclear, with the majority of experiments not reporting a response congruency effect. The review showed that the RTs in the studies that did not report an effect were slower than in the study showing an effect. This is in line with results in other tasks and is predicted by the leaky accumulator model. If the presence of a response congruency effect depends on the response speed then none of the reviewed theoretical accounts can predict this pattern. With respect to the experiments in the following chapters a response congruency effect is predicted in fast RTs, but not in slow responses.

1.5 Summary

Masked priming is a suitable method to shed light on early processing in visual word recognition. By keeping the presentation duration of the prime short the time available for stimulus processing is limited and hence effects from early stages become

apparent. The differences between findings from masked priming and experiments using longer presented and consciously available primes illustrate the different processes the methods tap into (e.g., Seguí & Grainger, 1990). It was shown that by varying the time allowed for prime processing, the source of priming effects differs (e.g., Ferrand & Grainger, 1992). Furthermore, masking can prevent effects of conscious strategy, though unconsciously presented primes can be exploited strategically, if the proportion of informative primes is unbalanced (e.g., Bodner & Masson, 2003).

Effects of response congruency between the implicit prime and the explicit target response were reported in various tasks (e.g., in number magnitude estimation, see Dehaene et al., 1998). In contrast, in lexical decision the empirical data and the predictions from theoretical accounts are less clear, e.g. the entry opening model does not predict a response congruency effect (Forster, 2004), but a deep processing account (Dehaene et al., 1998) predicts that a response congruency effect can emerge after applying the task instructions to the prime stimuli unconsciously. In empirical data, three studies have not reported an effect of response congruency (Norris & Kinoshita, 2008; Perea et al., 1998, 2010), but one study has reported an effect that cannot be attributed to other factors in the experiment (Jacobs et al., 1995). Also C. J. Davis and Lupker (2006) provided data that indicate a response congruency effect.

In order to resolve the theoretical dispute the empirical discrepancy needs to be resolved. Also, comparing those experiments that report an effect to those not reporting an effect, can help to reveal what information was processed in the very early stages of word recognition.

2. Orthographic typicality

This chapter reviews the experimental findings with regards to orthographic typicality and computational accounts for generating nonwords and selecting words based on sublexical properties. Concluding from the problems pointed out in the review, I introduce a measure for computing orthographic typicality. Finally, empirical evidence for the effectiveness of that measure is presented.

2.1 Review of orthographic typicality

Typicality can be defined as the result of a comparison of one entity to a whole set of entities with respect to a particular property. For example, Cassandra is known for her negative predictions (e.g., the destruction of Troy, the death of Agamemnon, her own demise). Another negative prediction like the abduction of Helena would be more typical than a positive prediction. In visual word recognition and in this chapter I refer to the typicality of a letter string with respect to its orthographic structure in comparison to the orthography of the whole language. A highly typical letter string resembles the orthographic properties of a language, e.g. *dound* in English. An atypical example would stand out of a text, because it is not like that language, e.g. *dqrki*. The definition of typicality is related to wordlikeness, but wordlikeness takes other properties into account as well. J. Humphreys (2008) showed that pseudohomophones like *brane* were rated more wordlike than other nonwords like *brone*. Pronounceable nonwords were also considered more wordlike than unpronounceable nonwords (also see Rubenstein, Lewis, & Rubenstein, 1971; Rubenstein, Richter, & Kay, 1975 for evidence in lexical decision). The rated wordlikeness was an effective predictor of RT in a speeded reading aloud experiment (J. Humphreys, 2008). In contrast, pseudohomophony and pronounceability do not directly influence the typicality of the orthographic structure. Though, indirectly a high similarity to an existing word could affect the typicality of a nonword.

2.1.1 Lexical frequency

Lexical frequency of words can be used as an approximation for the typicality of a word in a language, e.g. the high frequency word *the* is typical for English. The

logarithm of lexical frequency was shown to be a predictor of the required presentation time for reading aloud a tachistoscpically presented word (Howes & Solomon, 1951). Similarly, this measure predicted RTs in lexical decision (Whaley, 1978). This finding showed that words that are more frequent are recognised faster and with a higher accuracy than less frequent words. Though, the frequency measures collected in linguistic corpora, like SUBTLEXus (Brysbaert & New, 2009), provide an average for the usage in the population and thus, only an approximation for the experience of the individual participant. This was illustrated by participants showing differentiated frequency effects according to their personal interests and the associated vocabulary (e.g., Postman & Schneider, 1951; Solomon & Howes, 1951). The study of large databases with several thousand words showed that the averaged frequencies provide good predictions for responses averaged across participants (Brysbaert & New, 2009; Ferrand et al., 2010). Other measures than logarithmic frequency were also suggested, including the rank of lexical frequency (Murray & Forster, 2004, 2008; but see Adelman & Brown, 2008a, 2008b; Adelman, Brown, & Quesda, 2006) and experiential frequency (Gernsbacher, 1984). Furthermore, the effect of lexical frequency appears to be task dependent, e.g. it is attenuated in category verification tasks (Balota & Chumbley, 1984, 1990).

In general, a higher lexical frequency is robustly associated with faster responses and higher accuracy in lexical decision (Brysbaert & New, 2009; Ferrand et al., 2010). Thus, typicality as measured by the frequency of occurrence is an effective predictor for reaction times. In measuring the typicality of nonwords lexical frequency is not applicable. One option of using lexical elements in determining the typicality of nonwords is comparing them to words that are similar, e.g. in counting the number of neighbour words where the empirical evidence is reviewed in the next section.

2.1.2 Neighbours

In order to measure the typicality of a stimulus, this letter string could be compared to the most similar existing words. Coltheart et al. (1977) suggested counting the *neighbours* of a letter string, where a neighbour is a word that can be formed by replacing one letter of the string in question, i.e. all words with Hamming distance

(Hamming, 1950) of one from the target. The greater the number of neighbours *N* the higher is the typicality of that string. In Coltheart et al.'s (1977) lexical decision experiment an effect of *N* was found in the *no*-responses only. The nonword stimuli with a greater *N* received slower responses than those with a lower *N*. This indicated that nonwords that are more typical as approximated by *N* are harder to reject than less typical nonwords. This supported the idea of the *N*-metric. Following this study, a number of experiments investigated the effects of neighbours. Though Coltheart et al. (1977) did not report an effect in word targets, facilitatory effects of *N* were reported in low frequency words (Andrews, 1989, 1992, 1997) and inhibitory effects were reported if the neighbour was of a higher frequency than the target (Grainger, O'Regan, Jacobs, & Seguí, 1989).

Jacobs and Grainger (1992) showed that the semistochastic interactive activation model based on the original IA model (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982) can predict inhibitory effects of *N* in word targets as a result of competition simulated by lateral inhibition between word nodes. Using the dual-route cascaded model (Coltheart, Curtis, Atkins, & Haller, 1993) Coltheart and Rastle (1994) showed the facilitatory effect of *N* in words in their simulations. Their facilitatory effect was due to more word nodes receiving activation in high *N* words compared to low *N* words which lead to a higher top-down feedback to the letter units and in turn faster rising activity on the high *N* word detector. Crucially, the difference between these models is the impact of lateral inhibition (source of the inhibitory effect in Jacobs & Grainger, 1992) relative to the weight assigned to the facilitatory effects of activating a word detector embedded in a cluster (source of the facilitatory effect in Coltheart & Rastle, 1994). Grainger and Jacobs (1996) suggested including a parameter specifying the weight of the input of summed lexical activation by modifying the decision criteria of the model. This factor would depend on the specific task requirements, e.g. in an experiment where nonwords are not pronounceable and easily distinguished from the word targets the summed lexical activity was more beneficial and thus, a lower criterion was assigned. Similar results were obtained in the simulations in Chapter 8. A lower decision criterion for summed lexical activity (or a greater weight of summed lexical activity) would result in facilitatory effects of *N*, but a lower weight would

emphasise identification and competition between word nodes and thus, trigger inhibitory effects. This assumption is compatible with Andrews' (1997) claim that the different outcomes of the effect of *N* have to be attributed to different experimental contexts.

Relative frequency of neighbours

Another contribution to the different results of *N* effects is the relative frequency of the target and the neighbour. Neighbours of a higher frequency than the target are associated with inhibitory effects in word targets (Grainger et al., 1989) and facilitatory effects in nonwords (Grainger & Jacobs, 1996). That means that words with a higher frequency neighbour were responded to more slowly than words without such a competitor. Nonwords with a higher frequency neighbour were responded to faster, which could be attributed to faster identification of their relation to the neighbour (see Grainger, 2008 for review). Similarly, in measuring eye movements the existence of a higher frequent neighbour prolonged the gaze duration on target words (Grainger et al., 1989; Grainger, O'Regan, Jacobs, & Seguí, 1992). Though, in a similar experiment Sears, Hino, and Lupker (1995) reported a facilitatory effect of higher frequency neighbours in low frequency words. In masked primed lexical decision experiments, higher frequency neighbours interfered with the target response compared to a lower frequency neighbour (Grainger, 1990; Seguí & Grainger, 1990). Davis and Lupker (2006) were able to show facilitatory priming effects from neighbouring nonwords and inhibitory priming effects from word neighbours in one study. Davis (2003) attributed these findings to two components. First, all form-related primes produced activation in the target. This pre-activation allowed the quicker identification of the target. Secondly, (higher frequency) neighbours started to inhibit the target word node laterally. In sum these two processes resulted in inhibitory priming effects using word primes. The inhibitory component was emphasised in high frequency neighbours due to a higher resting level of these word nodes. Since nonwords are not represented lexically they do not produce lateral inhibition. Thus, the facilitatory effects in form-related nonword primes were attributed to the absence of lateral inhibition. The underlying mechanism is the same as explained above, the relative contribution of the benefit of summed activity in the lexicon and the disadvantage of lateral inhibition. In measuring typicality

the first component, summed lexical activity, is particularly interesting. Whilst inhibitory activity reflects local interaction between two word nodes, the summed lexical activity is a measure of how much the target letter string reflects the orthographic structure of the language.

Other types of neighbours

The most common definition of neighbours refers to letter strings with a Hamming-distance (Hamming, 1950) of one (Coltheart et al., 1977). The Hamming-distance allows only a like-for-like replacement (e.g., letter-for-letter, binary-for-binary, but not letter-for-space) but there are other operations for modifying a letter string. It could be argued the two strings *how* and *howl* are perceptually similarly distant as *able* and *axle*. The Levenshtein-distance is one alternative measure of editing distance (Levenshtein, 1966). It includes the substitution, deletion and insertion of letters. Each operation is counted and the original measure weighted them equally. Thus, the distances between *how*–*howl* and *able*–*axle* are both equal to one. Importantly, the length of the two stimuli is not fixed allowing more flexibility (see Davis, 2010 for discussion of length specific coding). Additional flexibility also arises from letter transpositions. Damerau (1964) considered the three operations above and transpositions and claimed 95% of typing mistakes could be explained by one of these four operations. Combining the two accounts results in the Damerau-Levenshtein-distance, which weights all four operations equally and measures the editing distance by counting the required steps to transform one string into another. That means that *drown*–*down*, *salt*–*slat* and *able*–*axle* all have the same distance of one operation. Different weights could be assigned to each operation for theoretical reasons, e.g. if one assumes that longer words inhibit shorter words but less so vice versa (see Davis, 2010), then shorter words are more confusable (i.e. more similar) to their longer neighbours and deletions were assigned a greater weight than additions.

The Damerau-Levenshtein-distance is a complex measure, but each single operation can be tested and reviewed in its own right by investigating the empirical effect of each single operation (transposition, deletion and addition; for substitution see the above discussion of *N*). Strong facilitatory effects were obtained using masked transposed

letter nonword primes in lexical decision (Lupker, Perea, & Davis, 2008 in consonant pairs only; Perea & Lupker, 2003a) and associative priming using lexical decision (Perea & Lupker, 2003b). Andrews (1996) found an effect of transposed letter neighbours in an unprimed lexical decision experiment, where low frequency words were facilitated by their higher frequency transposed letter neighbour and high frequency words were inhibited by their lower frequency neighbour. Higher frequency deletion neighbours interfered in a semantic categorisation task with the target response (Bowers, Davis, & Hanley, 2005). Deletion neighbours in nonwords and deletion neighbours with a higher frequency in words showed interference in a lexical decision task as well (Davis & Taft, 2005). In a masked primed lexical decision experiment facilitation from nonword primes derived from the target by deleting one letter (Schoonbaert & Grainger, 2004) and by adding one letter (Van Assche & Grainger, 2006) were reported. An inhibitory effect of addition and deletion neighbours was also found in eye movements and in lexical decision (Davis, Perea, & Acha, 2009). In summary, there is empirical evidence that each single operation transposition, insertion, deletion and substitution (see above) that is involved in computing the Damerau-Levenshtein-distance resulted in related letter strings with demonstrable effect on the recognition of the base word.

There is also evidence that the combination of two operations results in letter strings relevant for the neighbourhood of a word. C. J. Davis and Bowers (2004, 2006) conducted a series of illusionary word and masked primed lexical decision experiments showing that *N1R*s (*neighbour once removed:* a letter string formed by substituting and transposing a letter, e.g. *stop–soap*, *camp–clap*) and *DSN*s (*double substitution neighbour:* a string formed by replacing two letters, e.g. *snap–stop*, *camp–clip*) are similar to their base. Furthermore, C. J. Davis and Bowers (2004, 2006) showed that *N1R* strings have a greater similarity to the base word than *DSN*s. This suggests that in calculating the distance between two strings the weight assigned to transpositions should be lower than the weight assigned to substitutions. The effects of using related letter strings especially in masked priming were very successfully simulated by C. J. Davis (2010) using the Spatial Coding model.

Yarkoni, Balota, and Yap (2008) used the Levenshtein distance to compute the *orthographic Levenshtein distance 20*-measure (*OLD20*). This index is the average of the number of operations required to form the 20 closest neighbours of a word. The lower the *OLD20* value, the less operations were required to form the closest related strings. Thus, more wordlike letter strings are associated with lower values. They argued that including the transposition as a single operation instead of expressing it by two substitutions did not result in a marked difference and thus, used the three basic operations (insertion, deletion and substitution) with equal cost of one. This means that transpositions were assigned double the cost of a substitution which is contrary to C. J. Davis and Bowers' (2004, 2006) findings. Nonetheless, Yarkoni et al.'s (2008) regression analysis of the English Lexicon Project (Balota et al., 2007) showed that *OLD20* was the strongest predictor in speeded reading aloud out of a set of predictors comprising length, frequency and *N* in monomorphemic and in a combined set comprising mono- and multimorphemic words. Furthermore, the correlation of *OLD20* and lexical decision RT was almost as strong as of lexical frequency. Yarkoni et al. (2008) argued that *OLD20* reflects a measure of global similarity rather than a measure of local similarity like *N*. Computing a global measure of similarity like *OLD20* is different from computing the similarity of two specific strings (e.g., the match value in Davis, 2010). *OLD20* is rather computing the similarity between a string and a whole language (indexed by the 20 most similar items). This is compatible with the definition of typicality: one entity is compared to its domain of entities. Here, this means comparing a word or nonword to all words in a language. Thus, these findings can be interpreted as showing the effect of a global measure of typicality. Further empirical evidence of measures of global typicality are reviewed in the next section.

2.1.3 Global measures of typicality

The orthographic code is formed of letters, where the letters have a distinctive function but do not bear meaning themselves. Shannon (1948) regarded language as a stream of letter events. Looking at the frequency of each event there are more common and less common events. With respect to information density the less common letters are more informative and more distinctive. The informativeness of a letter is an important property in a noisy channel which was investigated by Shannon

(1948). It determines how informative a letter is for the whole sequence. That means letters that occur very frequently bear less information than rare events. They can also be guessed more easily, thus the term *redundancy* is used. This principle becomes apparent in looking at the number of four letter words containing an *e* (821/2431) versus the number containing an *x* (32/2431, Ramsden & Ramsden, 2007). The presence of an *x* constrains the set of potential candidates more than the presence of an *e* and thus, *x* is more informative (and less redundant) for identifying the string. In fact, the token frequency rather than the type frequency of each letter would have to be taken into account, but this example illustrates the principle of Shannon information (see Miller, Bruner, & Postman, 1954 for similar explanation). Furthermore, it shows the relation to the neighbourhood. In forming other words, it is more likely that the *x* has to be substituted than the *e*. Thus the *x* increases the average editing distance (e.g., OLD20 in Yarkoni et al., 2008). An increased editing distance also means that the likelihood of finding a substitution neighbour (*N*) is reduced.

Bigrams in tachistoscopic identification

An n-gram is an ordered tuple of n letters. A bigram refers to an ordered letter pair, e.g. the first bigram of *word* is *wo*, the second *or* and so on. The order of the letters is crucial in recognising words, e.g. in the pair *salt* and *slat* (see Davis, 2010 for discussion).

Miller et al. (1954) conducted a tachistoscopic identification experiment using nonword stimuli. They formed the stimuli by following the order of approximation outlined in Shannon (1951). Thus, the nonwords were random letters, letters generated according their frequency in English text, letter strings formed by using the bigram transition probabilities and finally quadrigram transition probabilities. Their results showed that across all presentation durations the number of correctly reported letters was greater the greater the order of language approximation. Miller et al. (1954) hypothesised that the amount of information processed by the participants was constant. Thus, they computed the information per letter and multiplied it with the correctly reported letters. Their results indicated that the processed information was not constant, but the recognition speed was dependent on the typicality of the stimuli.

In psychological research the frequency of bigrams received most attention as a measure of orthographic redundancy. Owsowitz (1963) explained that bigrams (or *digrams* in their publication) should be counted in a position specific fashion and the counts should only be used in words of the same length (also see Mayzner & Tresselt, 1962a, 1962b, 1963). In anagram solution tasks the positional bigram frequency (Mayzner & Tresselt, 1962a) and summed bigram frequency (referred to as word transition probability in Mayzner & Tresselt, 1963) showed a facilitatory effect on anagram solution time. Mayzner and Tresselt (1962b) asked participants to rank words according to their bigram frequency and found that participants were able to do so. Mayzner and Tresselt concluded that bigram frequency plays an important role in word recognition in addition to lexical frequency and word length. Owsowitz (1963) extended these findings and varied lexical frequency and summed positional bigram frequency in a tachistoscopic identification study with ten steps of luminance intensity. With regards to correct identifications, bigram frequency showed an inhibitory effect in both low and high frequency words. Also, words with high lexical frequency were generally identified more accurately. Owsowitz (1963) also investigated the letters participants entered in the report forms in low intensity conditions when they had not identified the word correctly. In these data bigram frequency showed a facilitatory effect on letters reported correctly, but lexical frequency did not show an effect. Owsowitz (1963) argued this reflects a guessing strategy because the more likely option is correct in high bigram frequency items. Alternatively, it was suggested that unusual letter combinations have higher threshold for being reported and thus, participants guessed a more likely letter combination. In more recent terms, the inhibitory effect of identification could be attributed to inhibitory effects from neighbours (see Andrews, 1997; Davis & Lupker, 2006; Seguí & Grainger, 1990) and the facilitatory effect on letter recognition to a successful guessing strategy (see Gernsbacher, 1984).

Interestingly, Biederman (1966) repeated the tachistoscopic experiment using the same 16 items as Owsowitz (1963). Biederman (1966) found that high bigram frequency items were identified faster and more accurately than low bigram frequency items. In a second experiment Biederman (1966) used better controlled stimuli with

regards to the distribution of bigram frequency within the word (see Mayzner & Tresselt, 1966; for a different perspective see Seidenberg & McClelland, 1989). The results were compatible with the advantage of high bigram frequency in the first experiment. Furthermore, the effect of lexical frequency was not found in high bigram frequency items and vice versa in high frequency items no effect of bigram frequency was found.

Broadbent and Gregory (1968) reported a similar tachistoscopic identification experiment. Their results resembled Owsowitz' (1963) findings where low frequency words were identified more accurately if they had low bigram frequencies. Broadbent & Gregory (1971) manipulated letter and word frequency in three tachistoscopic identification experiments. They did not report an effect of letter frequency nor of bigram frequency. But Broadbent and Gregory (1971) pointed out that a particular high frequent bigram occurred often in the experimental stimuli and thus, participants could have had to rely on the more rare parts of the word.

Rumelhart and Siple (1974) used a tachistoscopic identification task using all three letter words in Kučera and Francis (1967) and simulated their data as well as Broadbent and Gregory's (1968) data. Similar to Mayzner and Tresselt's (1966) and Biederman's (1966) findings, Rumelhart and Siple (1974) reported a facilitatory effect of bigram transition probabilities with better identification in high bigram frequency items. This is in contrast to other findings in tachistoscopic identification (Broadbent & Gregory, 1968; Owsowitz, 1963), but Rumelhart and Siple (1974) highlighted a difference in the experimental procedures. All experiments showing an inhibitory effect of bigram frequency used words only, whilst those reporting a facilitatory effect used both words and nonwords. Null effects of bigram frequency have also been reported. McClelland and J. C. Johnston (1977) observed no effect of bigram frequency and attributed the advantage of pronounceable over unpronounceable letter strings to structural properties. However, Massaro, Venezky, and Taylor (1979) pointed out that pronounceability and bigram frequency are correlated.

Rumelhart and Siple (1974) successfully simulated Broadbent and Gregory's (1968) inhibitory effects of bigram frequency by setting their P_{word} to 1.0. This means that the model could rely on the fact that the target was a word. Even though the effect on word identification was inhibitory, Owsowitz (1963) as well as Broadbent and Gregory (1968) reported a facilitatory effect of bigram frequency on letter identification. This indicated a general facilitatory effect of bigram frequency on the letter level. The effect of bigram frequency on the letter level was consistently facilitatory, whereas the direction of the effect on word identification seems to be task dependent. If a computational model of word recognition includes a parameter for capturing the experimental conditions, both inhibitory and facilitatory effects of bigram frequency can be successfully simulated (Rumelhart & Siple, 1974). The importance of including a parameter for capturing the experimental context was also highlighted by Grainger and Jacobs (1996) and the simulations in Chapter 8.

Bigram frequency in lexical decision

The effect of bigram frequency has also been investigated using the lexical decision paradigm. Rice and Robinson (1975) showed that the mean positional bigram frequency had an inhibitory effect in a lexical decision task in low frequency words. There was no such an effect in high frequency words and nonwords. A more pronounced effect in low frequency words compared to high frequency words resembles the effects typically found in manipulating the neighbourhood density, though an effect in nonwords would be expected (see 2.1.2). Jastrzembski (1981, Experiment 2) used the summed positional bigram frequency for measuring string typicality in a control experiment, but did not report an effect of this manipulation in their lexical decision experiment. Gernsbacher (1984) presented a review and replicated Rice and Robinson's (1975) findings. Though, the results were attributed to experiential frequency as rated in a separate experiment. The low frequency words in Rice and Robinson's (1975) stimuli set differed markedly on that variable between conditions and Gernsbacher (1984) attributed their finding to experiential frequency as well. In a second experiment, Gernsbacher (1984) used items of previous studies investigating bigram frequency (Biederman, 1966; Broadbent & Gregory, 1968; Owsowitz, 1963; Rice & Robinson, 1975; Rubenstein, Garfield, & Millikan, 1970) and

reported that earlier findings could be explained by experiential frequency. Furthermore, Gernsbacher (1984) showed in a factorial manipulation that experiential frequency but not bigram frequency was an effective predictor of RT.

Andrews (1992) provided a review and highlighted experimental findings of manipulating total positional bigram frequency and neighbourhood density. Andrews (1992) reported a facilitatory effect of bigram frequency in naming in low frequency words. In lexical decision, Andrews (1992) did not report an effect of neighbourhood density or bigram frequency and concluded that an effective manipulation involves both variables (also see Balota & Chumbley, 1984).

McClelland and J. C. Johnston (1977) attributed their results in tachistoscopic identification to the pronounceability of the letter string rather than to its bigram frequency. Using the lexical decision paradigm, a similar effect emerged: pronounceable nonwords received slower *no*-responses than unpronounceable nonwords (Rubenstein et al., 1971, 1975). This cannot be attributed to legality alone, since illegal unpronounceable nonwords received faster *no*-responses than illegal pronounceable nonwords (Rubenstein et al., 1971). Massaro, Venezky, and Taylor (1979) indicated that pronounceability and bigram frequency were correlated (Heyer, Quasthoff, & Wittig, 2006; also see Küpfmüller, 1949).

These results imply that the effects of bigram frequency, legality, pronounceability and most likely neighbourhood density are strongly related and could perhaps be summarised as wordlikeness. The size and the direction of the effect of wordlikeness depends on the context of the experiment (Andrews, 1992, 1997; Siakaluk, Sears, & Lupker, 2002), e.g. whether the experiment comprised words only or words and nonwords (Rumelhart & Siple, 1974). In general, effects of orthographic relatedness and surface structure are more apparent in low frequency words and nonwords than in high frequency words (Andrews, 1989, 1992, 1997; Gernsbacher, 1984).

With regards to high frequency words, Westbury and Buchanan (1999, 2002) reported results of lexical decision experiments where the probability of the least likely

bigram showed an effect in high frequency words. High frequency words with a low frequency bigram were responded to faster than high frequency words with a high frequency least likely bigram, i.e. without a low probability bigram. There was no such effect in low frequency words. These findings indicated a disadvantage of typical orthographic structure in high frequency words only and contrasted with previous findings where the effect of bigram frequency was limited to low frequency words. Along these lines, Jared (1997) showed a consistency effect in reading aloud in high frequency words. The effect of consistency was previously assumed to be most prevalent in low frequency words (e.g., Seidenberg & McClelland, 1989). But the results of Jared (1997) and Westbury and Buchanan (1999, 2002) suggested that wordlikeness of the orthographic structure can influence word recognition in high frequency words as well. Together with previously reported findings, this indicates that the orthographic structure can affect the recognition of nonwords and high and low frequency words.

For measuring orthographic typicality Westbury and Buchanan (1999, 2002) used the lowest bigram probability in a word where the lowest bigram probability corresponds to the highest density of information. Other researchers (e.g., Andrews, 1992; Gernsbacher, 1984) used some form of adding bigram frequencies or forming the average. The findings claimed as “paradoxical” by Westbury and Buchanan (1999, 2002) could be explained in less surprising terms. In high frequency words, it is likely that all bigrams are comparably high in frequency. Thus, a very low bigram probability that is paired with a comparable high level of information and distinctiveness could allow for faster identification by reducing the potential set of candidates. In contrast, low frequency words are more likely formed of low frequency bigrams. Thus, a low probability bigram is unlikely to stand out as it does in high frequency words. This could explain the absence of the effect of the least likely bigram in low frequency words. This explanation is based on the assumption that participants needed to identify the targets as in hard lexical decision tasks (e.g., Grainger & Jacobs, 1996), but Westbury and Buchanan (1999, 2002) did not provide a list of their stimuli. Specifically, the weight assigned to summed lexical activity is an important variable as explained above. The empirical evidence suggests that the average of bigram frequency could be more

informative in low frequency words, whereas a peak of information density (a very low frequency bigram) could be informative in high frequency words.

In a lexical decision study Hauk et al. (2006) compared typical (e.g., *yot*, *node*) and atypical stimuli (e.g., *yacht*, *gnode*) measured by position specific bigram and trigram frequency. In the behavioural data atypical pseudowords received faster *no*-responses than typical pseudowords. Additionally, Hauk et al. (2006) measured event-related potentials (ERP) where the effect of typicality was significant in word and nonword targets, even though behavioural data in words was inconclusive. The study employed very similar pseudohomophone-word pairs and provided evidence that the typicality of letter strings shows a significant effect in the very early stages of word recognition (approximately 100 ms after target onset) in the ERP data. There is also fMRI evidence that trigram frequency can influence the word recognition process. Binder et al. (2003) showed that nonwords with high mean positional bigram frequency activate brain areas that are more similar to those activated by words than those activated by low bigram frequency nonwords. Woollams et al. (2011) showed in a fMRI study, that orthographic typicality activated a brain area that was different from lexicality and concluded that there is an independent influence on the word recognition process. Westbury, Buchanan, Sanderson, Rhemtulla, and Phillips (2003) provided a genetic algorithm for computing significant effects with several interacting factors on neuroimaging data especially for the domain of lexical access. However, it is not as straight forward to extract cognitive conclusions from this evidence (Coltheart, 2006a, 2006b; Page, 2006). But Experiment 2 through to 9 showed that the response congruency and the typicality of the primes contributed to a lexical decision task. In particular, it was argued that more typical stimuli activated more word nodes and thus, the summed lexical activity was higher in typical than in atypical stimuli which then resulted in the observed effects in behavioural data.

In summary, manipulating the sum or average of bigram frequencies appears to influence low frequency rather than high frequency words (Andrews, 1992; Rice & Robinson, 1975). Measuring the lowest probability refers to a peak of Shannon information in the letter string and significant effects on RT were reported in high

frequency words (Westbury & Buchanan, 1999, 2002). The effect of orthographic redundancy may swap direction as the requirements of the experiment shift from identifying each word (a *hard* lexical decision task) to getting a reasonable idea whether the string is a word or not (an *easy* lexical decision task). This is compatible with Rumelhart and Siple's (1974; also see Grainger & Jacobs, 1996) suggestion in explaining the difference between an advantage of high bigram frequency in studies using word and nonword targets (Biederman, 1966; McClelland & Johnston, 1977) and a disadvantage in studies using word targets only (Broadbent & Gregory, 1968; Owsowitz, 1963).

2.2 Algorithms for generating nonwords

J. Humphreys (2008) asked participants to rate her word and nonword stimuli on a wordlikeness scale. The ratings were transformed to *z*-scores and provided a good predictor for naming and lexical decision latencies. Using a case alternation technique, wordlikeness as measured by the ratings was the only variable to survive this strong manipulation. This method was reliable in her experiments and could be used as a control routine like simulating the items (e.g., using the SCM, Davis, 2010). However, for generating large sets of nonwords computational methods that can give an estimate of the wordlikeness are required (Keuleers & Brysbaert, 2010).

There are various ways of computing and generating letter strings that have a certain level of orthographic properties. As outlined above bigram frequency can be used in a position specific and length specific way by summing, averaging or normalising it. Also, substituting or transposing letters in a word is an option for creating nonwords, but the typicality of the letter strings can vary amongst the results even if the base words are matched. As shown in the above, the typicality of the nonword foils is of crucial importance, not only for the presence or absence of effects in an experiment but also for the direction of these effects. The following section reviews previous computational accounts for generating nonwords and introduces the method used in this thesis in more detail.

2.2.1 WordGen

Duyck et al. (2004) introduced WordGen as a software for generating items. Duyck et al. (2004) reported that WordGen was able to produce 80% pronounceable nonword strings with appropriate settings, which forms a benchmark for other algorithms. The constraints for generating items available in the software included the initial and final position specific bigram frequency, the minimum position-unspecific bigram frequency (see Westbury & Buchanan, 1999, 2002) and the summed position-unspecific bigram frequency. Length, neighbourhood density *N* and lexical frequency for words served as bigram independent options. The software relied on the lemma databases of CELEX (Baayen, Piepenbrock, & van Rijn, 1995) for Dutch, German and English and Lexique (New, Pallier, Brysbaert, & Ferrand, 2004) for French. In generating word items, a random entry in the lexical database is selected and the respective wordlist is then parsed up to the first entry that satisfies the criteria. The nonword generation method generates a random string and then verifies the criteria. It is important to note that the underlying bigram frequencies were derived by adding the logarithmic frequency of respective words in the lemma database. The summed bigram frequencies of a generated item are inverted from the logarithmic values in the output. This software was not aimed at producing items of a certain degree of typicality, but rather at producing wordlike nonword stimuli. Duyck et al. (2004) particularly emphasised its use for research in German and Dutch as no such tool existed at the time.

Bigram frequency can be used as a tool for generating stimuli, but it does not allow a comparison across languages. This is due to different corpus sizes and a different distribution of words, e.g. in terms of entropy (Felici & Pal, 2008). Thus, the opportunity of generating and comparing stimuli across language was not filled by the software. Even though, Duyck et al. (2004) argued the usage of lemma frequency per million resolved the problem, structural differences between the languages could not be accommodated. For example, English has more monosyllabic words than German, and one might speculate that the total number of bigrams is more evenly distributed in English than in German. The reason is that if words are generally shorter then the bigrams must convey more distinctive information than in longer words. Thus, a bigram

that is low frequent (below average) in German might actually be just average in English. The difference in morphemic structure contributes to this problem, e.g. most words in German show inflectional morphology whilst this is not the case in English. This results in an inherent incomparability of the bigram frequencies across languages.

The WordGen algorithm relies on the lemma database in estimating bigram frequencies. In languages such as German most words in text show overt inflectional morphology, but the lemma database does not account for this. Thus, by relying on the lemma database the frequency of bigrams related to inflection is underestimated, whereas bigrams reflecting the stem forms are overestimated. This could result in overestimating the typicality of an item that resembles infinite morphology and underestimating the typicality of other perhaps more frequent forms. As a result the nonwords could reflect this truncated language rather than the actual language found in texts. For example, German verbs undergo an inflection for person and tense and nouns show plural and case. This means that forms coinciding with the stem form in the lemma database occur less frequent than the database suggests, whereas other forms are not listed at all. Another aspect of inflection is a change in word length, i.e. the infinite morpheme -*en* is longer than inflected forms such as -*e* and -*t*. As result, not only the bigram count in the lemma database but also the word length specific count is mislead. All German verb lemmata end in -*en*, but in a sentence -*en* is only seen in the first and third person plural forms of the present tense and infinitives. If most texts report about events and what has happened, e.g. newspapers, TV programmes or what the speaker has done, the most prevalent would be first (-*e*/-) and third (-*t*) person singular. That means the inflected forms are one letter shorter than the infinite form and thus, counting bigrams in a length specific way mistakes these words with a different length. The comparison of the CELEX GCT (German Corpus Type) raw data and the CELEX GOL (German Orthographic Lemma) processed data can highlight these differences. Due to noun plurals ending in -*en* which were summed in their singular stem forms in the lemma database, the total number of words ending -*en* is slightly smaller in GOL (814188 occurrences with a probability of 0.784 that -*en* is word final) than in GCT (925815 occurrences, 0.719 probability). The differences are more dramatic in -*e* and -*t*. In the GOL data words ending in -*e* occurred 327457 times,

equivalent to 0.065 probability that a words ends in -*e*. In the raw GCT data words ending in -*e* are more than twice as frequent with 893485 occurrences and it is also more likely that words end in -*e* with a probability of 0.157. The finding is similar in -*t*, with 398270 occurrences (probability of 0.080) in GOL and 601829 occurrences (probability of 0.106) in GCT. A similar argument holds for any other language where words undergo declination and conjugation processes with a morphological reflex, including Dutch.

Another problem is due to the implementation of the software. WordGen does not handle the special characters in languages other than English appropriately, including German the umlauts ä, ö and ü[1]. The user-interface allows to select "German" as language and one can enter a word like *Glück* [fortune] for checking its summed bigram frequency. The output correctly lists *GL*, *LÜ*, *ÜC* and *CK* as the bigrams forming the word *Glück*. Due to WordGen's inability of handling the *Ü*, the summed bigram frequency is computed as 2479, which is equal to the sum of the frequencies of *GL* (720) and *CK* (1759). This ignores two bigrams in a simple noun without producing any notice to the user.

2.2.2 LINGUA

LINGUA is the language-independent neighbourhood generator of the University of Alberta (Westbury, Hollis, & Shaoul, 2007). This software package was specifically designed to work with multiple languages and included a programme for building a corpus from text, e.g. text collections from internet resources. This enables the software to be used with any language of interest, even if there are no appropriate corpora available. Orthographic neighbourhood, lexical frequency in words and position specific and unspecific n-gram (e.g., bigram, trigram) frequency are measured according to the input corpus. The nonword generation process is based on a Markov chain and ensures that all n-grams occur in an existing word. Most importantly, the distribution of n-gram frequencies in the generated nonwords was said to resemble the features of the language in the input corpus based on either position specific or unspecific counts of the n-gram units. Westbury et al. (2007) did not provide any

1 Thanks to Vera Heyer (Universität Potsdam, Germany) for making me aware of this.

further details about the algorithm that is used in generating nonwords. The basic principle of a Markov chain is described in 2.3.

LINGUA generates pronounceable nonwords in basically any language though Westbury et al. (2007) did not provide a benchmark. Furthermore, LINGUA relies on the actual language data rather than a lemma database. Thus, it provided a solution to two problems observed with WordGen (Duyck et al., 2004). It can operate in any language and uses actual language data. LINGUA did not provide the option to generate nonwords that vary in typicality, e.g. if a set of atypical and typical nonwords was required for an experiment. Also, the actual properties of the generated nonword strings are not accessible to the user and the computational mechanism was not clarified. Hence, the software provides a good possibility of generating typical and pronounceable nonwords, but the user does not have the option of manipulating the output in a controlled way. Similar to WordGen, LINGUA does not produce a language independent measure of wordlikeness or typicality.

2.2.3 Wuggy

Wuggy (Keuleers & Brysbaert, 2010) generates nonwords based on the syllabic structure of a language. The programme requires a corpus of orthographically syllabified words and a list of all possible orthographic syllable nuclei. This list of nuclei is essential for the software, but it can depend on the specific account of phonology. Keuleers and Brysbaert (2010) classified [ʊ] as a consonant in diphthongs such as [a͡ʊ] and thus, concluded [ʊ] is part of the coda (for other accounts see e.g., Hall, 2000; International Phonetic Association, 1999; Ladefoged & Johnston, 2011; Pompino-Marschall, 2009). Nonetheless, Wuggy appeared to use the longest possible orthographic nucleus when generating nonwords, e.g. *ou* in *house* despite Keuleers and Brysbaert's (2010) argument that [ʊ] and thus, *u* was part of the coda. The reason for *ou* appearing in the list of nuclei was *touch* where *ou* [ʌ] formed the nucleus. Thus, Wuggy avoids problems in nonwords that may have arisen from Keuleers and Brysbaert's (2010) account of phonetics.

Using the syllabified input corpus and a list of possible nuclei, all words and most nonwords (by using a heuristic) can be divided in onset, nucleus and coda straight forwardly. Wuggy measures the bigram frequency between two syllabic elements, i.e. onset and nucleus or nucleus and coda, by the lexical frequency of the respective words. The default nonword generation process relies on substituting elements in a base word. The algorithm keeps the orthographic length of each element (onset, nucleus, coda) constant per default and the number of substitutions is equal to the number of syllables in the base word. A so called *concentric search* finds the element that will be substituted. The algorithm searches for bigrams with a specific length ([word beginning, onset], [onset, nucleus], [nucleus, coda], [coda, word end]) that have a similar frequency as the corresponding bigram in the base word. If there are no suitable candidates within the initial acceptable range of deviation from the original frequency, the acceptable deviation is increased step by step. This forms the concentric search around the base word. Keuleers and Brysbaert (2010) claimed Wuggy produces the most similar word or the most similar nonword with respect to the base word (but see Davis, 2010 for a review of string similarity).

By using the most similar nonwords to a set of words an additional difficulty could arise in a lexical decision experiment. The participants could be confronted with nonword foils that are matched for wordlikeness with the word targets, but that are also hard to distinguish from their very similar word counterparts. This contrasts with normal reading where it is more likely that the context of other words is supportive or neutral (e.g., Fischler & Bloom, 1979; Inhoff, 1984; Schuberth & Eimas, 1977). A situation where the context is particularly not helpful occurs in a written tongue twister, e.g. *The sixth sick sheikh's sixth sheep's sick*. An experiment with highly similar word and nonword targets is likely to incorporate a difficulty from distinguishing the stimuli, i.e. responding *no* when a very similar word was presented a short time ago. This difficulty is on top of the difficulty of distinguishing words and nonwords and is independent of the wordlikeness of the stimuli. Since the confusability of targets appears out of direct control of the experimenter it should be avoided whenever possible. Keuleers and Brysbaert (2010) argued that the user could provide another set

of words that the nonwords would then be fitted to, but finding additional appropriate word targets could be another difficult to achieve enterprise.

As pointed out with other algorithms, Wuggy does not produce a language independent measure of wordlikeness or typicality. The user has to rely on the algorithm in Wuggy in finding the most similar nonword or word string to a given base word.

2.2.4 Summary

The available software relies on n-gram frequencies, either by summing the frequencies of letter bigrams (WordGen: Duyck et al., 2004) or bigrams of subsyllabic units (Wuggy: Keuleers & Brysbaert, 2010). Only LINGUA (Westbury et al., 2007) used a Markov chain for estimating the typicality of the stimuli, but no further details were provided. A problem with all programmes is that there is no language-independent indicator of how typical the generated items are. With regards to language independent functioning, LINGUA provided the most comprehensive tool-kit and could be used with any language. WordGen relies on preassembled corpora in the programme (LEXIQUE and CELEX) and Wuggy requires a syllabified corpus (most dictionaries provide a syllabification, but the list of nuclei could be difficult to find). In generating nonwords, the programmes aimed for mimicking either a particular word or forming nonwords that are coherent with the language average. But for the experiments introduced in this thesis it was necessary to specifically vary the nonwords on the wordlikeness scale, including extraordinarily wordlike nonwords on the one extreme, but also very atypical nonwords on the other end of the scale. The software that was used for this purpose is introduced in the next section.

2.3 Orthographic trigram transition typicality (OT3)

The following section introduces software for analysing already selected words and nonwords as well as for selecting words and generating nonwords. It enables the user to specify how wordlike the generated nonwords should be. Following a discussion of the underlying algorithm and the differences between the Orthographic trigram

transition typicality (*OT3*) and other measures, experimental evidence that this algorithm effectively generates nonwords of either high or low typicality is provided.

2.3.1 Differences to other algorithms

In the following section, I introduce the main differences between the *OT3* algorithm and measures of orthographic typicality that have been used in the literature. The main differences between the standard word length and position specific summed bigram frequencies and the *OT3* metric are the use of transition probabilities instead of frequency, the word length and position unspecific method of counting, the use of trigrams instead of bigrams and finally the use of a product instead of a sum.

Frequency and transition probabilities

The terms transition probability and frequency have sometimes been used synonymously in the literature, specifically in orthographic redundancy (e.g., Keuleers & Brysbaert, 2010; Mayzner & Tresselt, 1962a). However, there is an important difference between the two measures with respect to scaling.

The frequency of a bigram is the number of its occurrences in a corpus. Often the frequency is scaled by the number of words in the corpus and is presented as the number of occurrences per million words (e.g., Davis, 2005). This scales the number of letters to the number of words and makes data in English comparable across corpora. The comparability does not hold across languages, where a language with a more overt morphology than English could show a tendency to have less words to express the same content and as a result have tendency for more longer words. For example, the distinction between definite reference (*the*) and indefinite reference (*a*) could be expressed morphologically, making the use of the equivalent to the most frequent English word (*the*) redundant which is the case in most Slavic languages (Krifka, 1992).

In contrast to frequency, transition probabilities are scaled towards the local context and thus, are robust with regards to the average length of words or the corpus size (as long as the corpus appropriately reflects the language). In general, a transition

probability quantifies how likely it is that one event follows another. In bigrams, a transition probability describes how likely it is that the second letter of the bigram follows the first. It is computed by dividing the frequency of a bigram (e.g., *ab*) by the frequency of the first letter in that position (i.e., the frequency of *a* as a first letter in a bigram). This results in the probability that an *a* is followed by *b* (and forming the bigram *ab*). In other words, it is the probability of *b* under the condition that *a* is already known to be true, which reflects Bayes' theorem (Bayes & Price, 1763; Heyer et al., 2006; Küpfmüller, 1949). Each transition probability is the reciprocal of the Shannon information for the ordered relation between the letters, i.e. the information that *a* is followed by *b*. The higher the transition probability the less informative is the transition. The less informative a transition is (higher transition probability) the greater is the redundancy within the letter string.

The examples in Table 2.1 show that even though the frequency of bigrams in the word final position is very similar, their transition probability can be very different and vice versa. The first example illustrates that almost all five letter words (98%) that show an *h* in the fourth position are actually words that end in *ht*, whereas only 15% that show an *e* in the fourth position end in *ed*. In contrast, the frequency of the bigrams *ht* and *ed* is comparable and does not reflect the specific high orthographic redundancy of words ending in *ght*. The second example (of *...nd* and *...gh*) shows that the frequency of a bigram can also differ even if the transition probabilities are almost equal. The comparable transition probabilities show that the proportion of continuations is similar, i.e. the coherence of the pattern is similar, even though the frequency of words

Table 2.1: Position and word length specific bigram frequency and transition probability in five letter words according to N-Watch (C. J. Davis, 2005).

Bigram	Examples	Frequency/ million words	Transition probability
...ht	*eight, fight, right*	2727	0.98
...ed	*fried, speed, taxed*	2739	0.15
...nd	*blind, pound, spend*	1493	0.14
...gh	*cough, dough, laugh*	183	0.14

In all examples "." is used as a wildcard character to replace any single letter.

showing this pattern is different. It could be argued that for measuring the redundancy within the orthographic code, the transition probability is more appropriate.

Position and length dependent measures

Most bigram measures were provided in a length and position specific way (Davis, 2005; Mayzner & Tresselt, 1962a, 1962b, 1963; Owsowitz, 1963). But measuring the orthographic redundancy in words of the same length implies that there are different, unrelated orthographies for each word length. Furthermore, this account implies that these word length specific orthographies would not interfere with each other. The result would be that the *st* in *street* is unrelated to the *st* in *streets*. C. J. Davis (1999, 2010) argued that this is an implausible assumption. Furthermore, it was argued that position specific coding of letter positions is implausible and empirically not justified (e.g., Davis, 2010). Thus, it could be argued that counting bigram frequency or transition probabilities in a position and word length specific way is similarly implausible.

For example, the bigram *ld* is relatively frequent as the fourth bigram in five letter words and in the fifth position in six letter words, whereas it never occurs in the first position of five letter words and six letter words (see Table 2.2). These data indicate

Table 2.2: Position and word length specific bigram frequency and transition probability of *LD* in five and six letter words according to N-Watch (C. J. Davis, 2005).

Bigram	Examples	Frequency/ million words	Transition probability
ld...	NIL	0	0.00
.ld..	*elder, older, oldie*	116	0.02
..ld.	*colds, moldy, tilde*	43	0.02
...ld	*child, world, yield*	6250	0.51
ld.....	NIL	0	0.00
.ld...	*elders, eldest, oldish*	36	0.01
..ld..	*folder, golden, seldom*	167	0.05
...ld.	*fields, mouldy, yields*	100	0.03
....ld	*behold, shield, should*	1023	0.10

In all examples "." is used as a wildcard character to replace any single letter.

two patterns. First, it could be assumed the bigram never occurs in word initial position and secondly, it appears to be relatively frequent in word final position. Due to the length specific metric it cannot be said for sure whether the bigram never occurs in initial position until data of all word lengths was checked. The expression word final is only defined for specific word length in a length specific measure. In fact, from the five letter word data one could assume the bigram is frequent in position four, but this will be disappointed in six letter words. Thus, a metric that allows capturing these properties of an orthography should be word length and position unspecific.

Bigrams and trigrams

There is more information in a metric that relies on trigrams than in bigrams. Thus, it could be argued that the bigram measures introduced above are less complex than a metric of trigram transition probabilities and should be preferred on these grounds. But it is important to note that the word length and position specific bigram frequency is a very rich measure, because there is not only information about the directly preceding letter (as the name *bigram* might suggest), but also there is information about the exact position in the word. That means the word length and position specific bigram transition probabilities as they were computed in Table 2.1 and 2.2 do not resemble the probability of one letter following another, but they reflect information on the number of preceding letters, the exact position in the word and the number of following letters. For example, there is no obvious reason why an *l* should not follow the letter *b*. Words with the bigram *bl* include *blue, able, oblige, cable, edible* and so on. In a word length and position specific bigram measure, the frequency of *bl* drops to 0 in second and fourth position of five letter words. With regards to the fourth position the measure takes into account that this is the final position in the word even though this is not obvious from the bigram *bl*. The drop in the second position has a less obvious reason as there are examples of words with a *bl* in this position (e.g., *able*, *oblige*). It is simply that there is only one word with *bl* in that position and a word length of five letters (abler, Ramsden & Ramsden, 2007) and this is so low in frequency that a zero is displayed. That means the metric does not capture a regularity or a constraint on information density, rather this particular case is the result of an overly specific algorithm. Overspecification means the algorithm is not sensitive to rules or

generalisations, but rather captures singularities. In summary, a word length and position specific bigram measure is more rich in information than just the frequency of that bigram. In the following, I will refer to the actual bigram frequency, which is the frequency of a particular bigram counted according to the number of its occurrences irrespective of the word length and position in the word.

Küpfmüller (1949) investigated the redundancy of natural language transmitted through a noisy channel. By manipulating the length of n-grams Küpfmüller (1949) showed that trigrams produce very language like letter strings. In contrast, the smaller bigram and unigram (single letters) units tended to result in gibberish. Larger units such as quadrigrams tended to produce words which is a sign of overspecification. Producing words is very language like, but it does not allow for new input and in the context of information transmission utterances could be overly corrected (see Heyer et al., 2006 for similar argument). When generating experimental stimuli it would be impossible or at least very unlikely to generate nonwords. Manning and Schütze (1999) used n-grams on word level in generating language. They aimed for generating new but language like (grammatically correct) sentences. The optimum in avoiding plain reproductions of the input and producing language like utterances was trigrams which is compatible with the results at a letter level (Küpfmüller, 1949).

Table 2.3: Bigram and trigram frequency and transition probability according to BNC (*The British National Corpus*, 2001) for the letter string *STRST*.

Bigram	Frequency	Transition probability
#s	59765	0.07
st	30631	0.13
tr	10774	0.03
rs	12665	0.05
st	30631	0.13
t#	84269	0.23
Trigram	**Frequency**	**Transition probability**
#s	59765	0.07
#st	9517	0.16
str	3179	0.10
trs	0	0.00
rst	1559	0.13
st#	10949	0.36
t#	84269	0.10

Please note, the '#' replaces the space character. The transition probability for the bigram 't#' is based on different metrics. In the bigram, section this probability reflects the chances that a 't' is succeeded by the word end (or space character), thus the denominator is the number of 't' in the corpus. In the trigram section, this probability refers to the proportion of words ending in 't'. That means the denominator is the number of all words.

In correcting for orthographic errors, trigrams can be more useful than bigrams. For example, a typo like *jugde* (example from Perea & Lupker, 2003a, 2003b) could be corrected by software finding that *ugd* never occurs in English (using CELEX, Baayen et al., 1995). This makes the transition from *g* to *d* a likely typo. Using Damerau's (1964) basic operations the guesses would be *judge* (transposition), *juge*/*jude* (deletion), *jug.de* (addition, where . denotes any single letter) and *ju.de*/*jug.e* (substitution). Using trigram transition probabilities *judge* (transposition) is the most likely candidate, followed by *jude* (deletion) and *judde* (substitution). Correcting this mistake on the basis of bigrams would be more challenging because all bigrams in the typo occur in the English CELEX database: *ju* (*juice*), *ug* (*sugar*), *gd* (*kingdom*), *de* (*mode*). This example shows that bigrams cannot capture regularities of letter clusters, because two

letters already form the horizon of observation in bigrams. The example in Table 2.3 shows the bigram and trigram metrics for the letter string *strst*. This nonword is clearly unpronounceable in an English context. But pronounceable strings can start in *str* (*street*) and end in *rst* (*first*), also the combination of *rs* (*nurse*) is perfectly acceptable. In a left to right parse the letter string becomes unpronounceable as soon as no vowel follows the initial *str.* As Table 2.3 shows, a trigram metric can capture this fact, but the bigram metrics do not reveal any reason to doubt the legality of *strst*. Indeed, according to bigrams this letter string should be a rather typical example of the English language.

Sum and product

It was argued above that the use of transition probabilities is preferable to the use of frequencies. By the use of transition probabilities each n-gram can be regarded as an event. A word is then constituted by the simultaneous occurrence of these events. For example, all n-grams involved in the representation of *judge* have to occur together in order to form the word *judge*. The occurrence of a single event, e.g., the trigram *udg*, does not constitute the word *judge*. From the statistics this constitutes an AND-relation between all the events involved, thus the probabilities of the individual events have to be multiplied to accurately represent this combination of events. In psycholinguistics the logarithm is often used, where the sum of logarithmic values is equivalent to the product. It is important to note that frequencies and transition probabilities reflect a different normalisation and thus, these numbers are not equivalent irrespective of the mathematical operations applied. That means for forming the mathematically exact account of the orthographic typicality using transition probabilities, they should either be multiplied or the logged probabilities should be added.

Forming the product of transition probabilities and forming the arithmetic mean of transition probabilities is not only mathematically different, the predictions differ as well. In Table 2.3 it can easily be seen that the product of the trigram transition probabilities is equal to 0, whereas the mean is equal to 0.13 indicating a rather typical item. In this example, the product produces the desired result. The use of the product and the sum of transition probabilities was also compared empirically. Loth (2007)

compared both methods in predicting RT in a lexical decision experiment. The results showed a larger effect of orthographic typicality when the product was used in the computation compared to the arithmetic mean (which is the sum divided by a constant). This was attributed to a more reliable measure when using the product of transition probabilities.

2.3.2 *OT3* Algorithm

The aim of the *OT3* algorithm is to provide a measure of how likely a letter string is compared to the orthographic structure of a given language model (the corpus). The algorithm provides an estimate of orthographic typicality. This differs from wordlikeness, which includes other additional factors (Grainger & Jacobs, 1996; J. Humphreys, 2008). For example, J. Humphreys (2008) showed that pseudohomophones are more wordlike than other nonwords, but pseudohomophony does not affect the typicality of that nonword per se. Indirectly, the existing word would contribute to the transition probabilities and thus, influence the estimate of typicality. Similarly, pronounceability is not taken into account directly. Pronounceable nonwords are usually considered more wordlike (J. Humphreys, 2008; Rubenstein et al., 1975), but their pronounceability as a categorical distinction does not change the typicality of a nonword even though pronounceability and typicality are well correlated (Massaro et al., 1979). In the following section I introduce the implementation of the *OT3* algorithm.

The algorithm was implemented in JAVA (*Java Runtime Environment*, 2011) and operates platform independently with a JAVA Runtime environment. The linguistic input for the software is a corpus that specifies word frequencies which can be provided in actual frequency, frequency per million, the log of any frequency measure or any other numerical value. The alphabet is another required piece of information and is provided in a file listing all valid letters of the language under consideration. This specifies the letters that will be used in generating items, but it also specifies a filter for the corpus where items with letters that are not part of the alphabet are ignored. These items are not included in measuring transition probabilities.

2.3.3 Initialising

The programme provides an *Initialise* tab, where the user specifies the folder containing the *wordlist.txt* and *alphabet.txt*. This folder also serves for saving the computed transition probabilities and a list of excluded words. Both can be used to check the functioning of the algorithm.

The initialisation process parses the wordlist and computes the transition probabilities of all trigrams. In order to start a Markov-chain, the algorithm has to use a unigram in the first step, a bigram in the second step and trigrams for all further steps. In symmetry to the start of the chain, the end is formed by a bigram and a unigram.

The first and the last element in a word is always the start or end symbol and its probability is always 1.0. Thus, the two unigrams can be safely ignored. The bigram refers to the combination of word beginning (*#*) and the first letter (*a*). Thus, it is equal to the probability that a word starts with a certain letter (*#a*). This transition probability is computed by summing the frequencies of all words starting with a specific letter and dividing these counts by the total word count. As a result of this normalisation, it does not matter whether frequencies are provided per million or as total values as mentioned above. For finalising the chain, the bigram transition probability of the last letter and the word end is computed in the same way. Note, that using an additional space character and using trigrams from start to end is only equivalent at the beginning of word (the probability that a word starts with a certain letter), but not at the end of the word. The space character would have a probability of 1.0, because it is the only option to continue once the first space character (*bc#*) was encountered (leading to *c##*). This would not reflect the probability of word ending in a certain letter (*c#*) and result in an asymmetric Markov-chain.

The trigram transition probabilities are computed by adding the frequencies of all words containing a particular trigram (*abc*) and dividing this by the total number of occurrences of the corresponding initial two letters (*ab*). That means the first trigram in each word contains a single space character (*#ab*) and so does the last trigram (*bc#*). The trigram transition probability is equivalent to the normalised frequency of a trigram and thus, scales the provided frequencies. This also implies an important property of transition probabilities: their dependence on other alternatives. If *abc* has a frequency of 40 and it is the only relatively frequent continuation of *ab*, the transition probability will be high. In contrast, there could be other options (e.g., *aba* in *abacus*) with a much higher frequency and in that case the transition probability will be lower. This is illustrated in Table 2.4 with *gre* and *ree*. Despite their similar frequency, the transition probability of *gre* is more than eight times greater then for *ree*. Thus, the typicality of each transition depends on the local set of alternatives. This resembles the idea of local competition in computational models (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982). A medium frequency word node was the top candidate in a cluster of very low frequency competitors, whereas another node with the same frequency in a cluster of very high frequency word nodes would make the medium frequency word a less competitive candidate. In transition probabilities, the other potential candidates scale the probability of each transition. This marks an important difference to using the raw frequency data.

Table 2.4: Examples *event* and *green*. The table shows the frequency of the respective n-gram as in the English Corpus Type database of CELEX and the transition probabilities P_{trans}.

Trigram	Frequency	P_{trans}	Trigram	Frequency	P_{trans}
#e	407834	0.023	*#g*	330073	0.019
#ev	68166	0.167	*#gr*	73723	0.223
eve	134912	0.818	*gre*	41178	0.382
ven	58052	0.108	*ree*	49172	0.046
ent	266334	0.316	*een*	82714	0.289
nt#	213006	0.394	*en#*	296153	0.352
t#	1792867	0.102	*n#*	1444738	0.082
log product of transition probabilities		-5.36			-6.207
OT3 (standardised product)		1.64			1.225

Computing the initial and final transition probabilities is including position specific information in the chain. That means the initial and final trigrams and all bigrams include information about their position at the edge of the letter string. There are three reasons for including this information in processing the typicality. First, there are empirical reasons for treating the first and the last letter of a letter string special. The exterior letters can be reported more accurately than interior letters (Merikle, Coltheart, & Lowe, 1971; Mewhort & Campbell, 1978). Furthermore, including these letters in a transposition (Perea & Lupker, 2003b; White, Johnson, Liversedge, & Rayner, 2008) or replacing them (Schoonbaert & Grainger, 2004) is more disruptive to the recognition than manipulating internal letters only. The Spatial Coding Model (Davis, 2010) that was used in the simulations in Chapter 4 and 8 also treats these letters specially. Also, ignoring the special position of first and last letters would imply text comes without spaces, which is clearly not the case. Separating words by spaces facilitates reading (Grainger & Jacobs, 1999). Secondly, the initial letter marks the beginning of a syllable and the final letter the end of a syllable. This is important phonological information for starting and ending the generating process in a grammatical way. A specific marking of the position in the phonological stream is not required in the interior letters as they have an existent preceding and succeeding item to convey this information. Finally, the Markov-chain requires a start and an end. By starting with a trigram that covers the first two letters, the information about their typicality would be lost. Thus, it could be argued treating the begin and the end of a word special is justified. But it is important to note, that this positional information is still minimal compared to a word length and position specific account.

The output of the initialisation procedure are five files. First, the *ini_ok.csv* contains a line by line copy of all words that were accepted by the alphabet filter rule and *ini_nk.csv* the items rejected by the filter. Third, the file *bigrams.csv* contains a list of all possible first and last letters with their frequency and the respective probability. Fourth, the *trigrams.csv* contains a list of possible trigrams as specified by the alphabet and their respective frequencies and transition probabilities. Finally, *ini_standardisation.csv* contains the data (mean and standard deviation for each word

length) that is required for computing standardised values of typicality which is subject of the next section.

2.3.4 Analysing words

In analysing words the algorithm relies on a Markov-chain model (see e.g., Heyer et al., 2006). The chain is defined by the letter string that is analysed, i.e. the letters in the string define which transition probability has to go into the computation. In analysing letter strings both requirements of a Markov chain are satisfied, the chain has a defined start and end and there is exactly one path in the model. That means that the analysis is a finite state automaton where every analysis of a specific string produces the same result. In order to ensure this, the probabilities are fixed once they have been established during the initialisation.

A letter string is defined by the co-occurrence of all its letters in a particular order. In order to describe the co-occurrence of events, the product of the transition probabilities is computed as described above. Because transition probabilities range between 1.0 and 0.0, each multiplication reduces the magnitude of the result. The first consequence is that the resulting numbers are very small and in order to scale them the logarithm of the product is used (see examples in Table 2.4). This log product expresses the magnitude of the likelihood of an item with respect to a language (as defined by the input corpus). By using the logarithm, the distribution of values is also transformed. In particular, the differences between very low values (very atypical items) is reduced compared to the more typical items, because there are more atypical items than typical ones. This can be explained by the amount of random letter strings compared to the number of actually existing words. For example, there are 11.9 million possible combinations of five letters but only about 4500 five letter words (Ramsden & Ramsden, 2007). Thus, the small differences between very atypical letter strings are less important than between two typical items. A second consequence is that the measure is dependent on the length of stimulus. Even if strings are formed of equal letters, e.g. *xxxxx* and *xxxxxx*, the additional multiplication with a transition probability shifts the log product by one magnitude. This is despite the fact that five or six *x* in a row are similarly unlikely in forming a word.

A standardisation procedure is required for forming a measure that is independent of the length of an item. As a side effect, this standardisation provides a language independent measure for typicality. The *OT3* algorithm continues with calculating the log product of each word in the corpus. A length specific routine calculates the mean and the standard deviation of the log products of all words. For saving computation time, this is done once during the initialisation process and a table specifying the values is saved on the hard disk. Using these numbers a *z*-score value of typicality of each stimulus can be computed, where a high value refers to highly typical stimulus and a low value to an atypical stimulus. By expressing typicality in terms of standard deviations, the measure becomes more intuitive and language and length independent. This standardised log product will be referred to as *OT3* value.

For illustrating the standardisation process all 5, 6, 7 and 8 letter words of the English Corpus Type database of CELEX (Baayen et al., 1995) were plotted with their log product of transition probabilities in Figure 2.1. Each word length forms a different layer which is due to an additional multiplication for each additional letter. Figure 2.2 shows the respective *OT3* values (standardised log product). The shape of the standardised curves is the same as in the unstandardised case. Thus, if stimuli of only one length in one language are used the standardisation procedure has no effect, apart from shifting the numbers. The *OT3* plots for each letter length match almost perfectly indicating that the standardisation procedure showed the desired effect. Interestingly, the distribution of typical and atypical words is similar to the typical RT distribution.

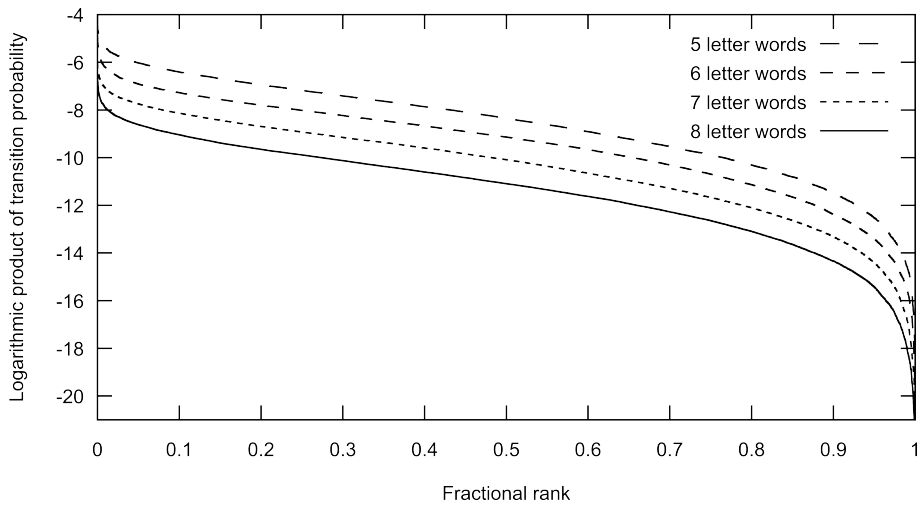

Figure 2.1: Plot of the log product of transition probability of 5, 6, 7 and 8 letter words in the English Corpus Type database of CELEX. Note, that the fractional rank was used. This allows a comparison of the spectrum of values even though the number of words per length differs.

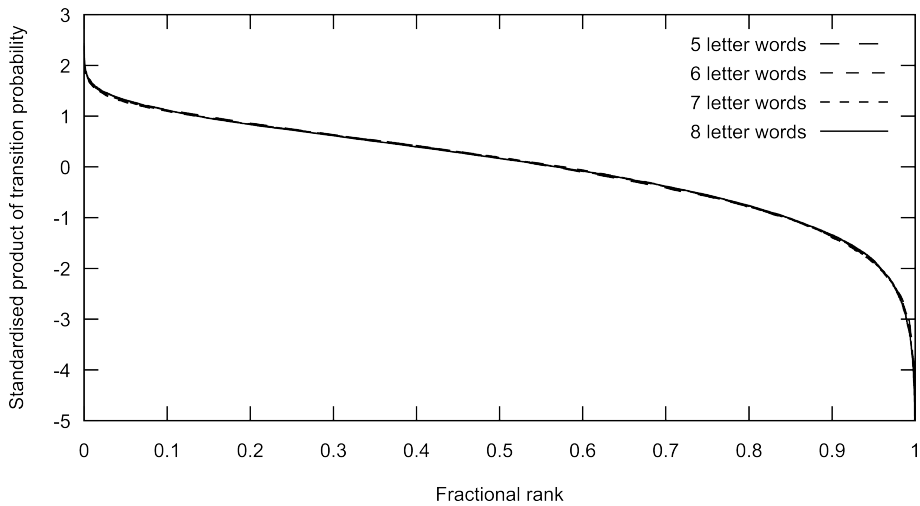

Figure 2.2: Plot of the standardised product of transition probability (OT3) of 5, 6, 7 and 8 letter words in the English Corpus Type database of CELEX. Note, that the fractional rank was used. This allows a comparison of the spectrum of values even though the number of words per length differs.

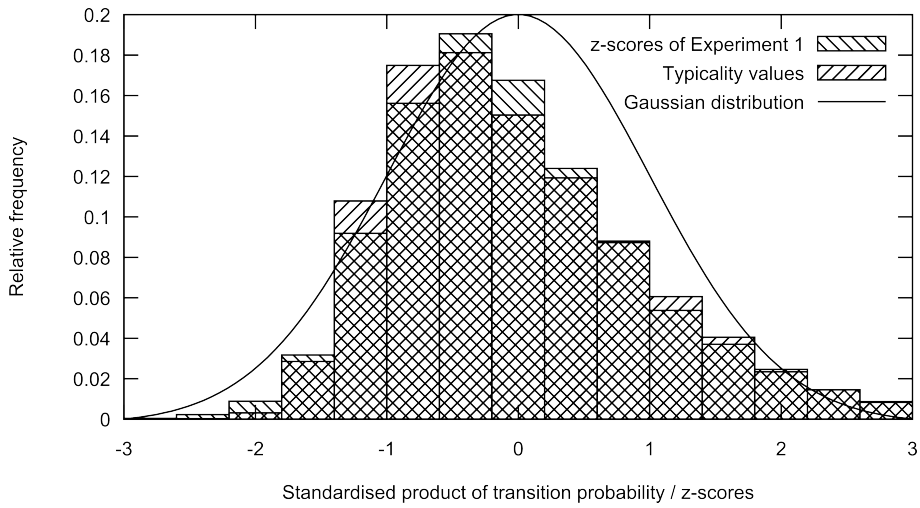

Figure 2.3: Histogram of OT3 values of all 5, 6, 7 and 8 letter words in the English Corpus Type database in CELEX and z-scores of RT of Experiment 1. Please note that the OT3 values were multiplied with -1 for plotting them in the same direction as the RT distribution.

There are fewer very typical than atypical words. Figure 2.3 shows the comparison of all 5, 6, 7 and 8 letter words in the ECT database of CELEX and the *z*-scores of RT of Experiment 2. Both distributions are very similar giving rise to the idea that the method applied for analysing word stimuli results in a natural distribution. But note that the similarity of the distributions does not imply any correlations between RT and typicality.

2.3.5 Generating items

The software provides a panel for generating items where the user has to specify the folder with all the files that were produced during the initialisation process. The user specifies the length of the stimuli and whether they should be words or nonwords. Also, the software provides two search modes, a random walk and an exhaustive search. The random walk is essentially an heuristic that produces random letter strings of a certain length which are then compared against the criteria specified by the user. Those items matching the criteria are sent to the output and the procedure continues until the number of items specified by the user is hit. In an exhaustive search, the algorithm parses through every possible letter combination. For reducing the computational load the least demanding metric is computed first for excluding items from the output. That means the *OT3* value is essentially a look-up in a small dataset of transition probabilities and can be computed faster than *N*, which requires

numerous look-ups in the lexicon database. If the maximum number of stimuli is expected to be small, an exhaustive should be performed. Otherwise, the programme would continue listing a few items repeatedly until the specified number of items is reached. Where a large number of possible items is expected, the random search provides the user with the desired number of stimuli.

There are several optional values that can be specified by the user, these include the maximum and minimum number of neighbours. Additionally, the arithmetic mean of transition probabilities can be specified. This option also allows to specify a maximum deviation of the transition probabilities in the string. By using this option, it can be specified how homogeneous the transitions within a stimulus are. For example, a nonword could have a peak of a very unlikely trigram embedded in typical letter combinations. In bigram frequency this was referred as the consistency within the string (Biederman, 1966; Mayzner & Tresselt, 1966), whereas strong deviation marking other boundaries (e.g., morphological, syllabic) were referred as bigram troughs (Seidenberg & McClelland, 1989).

Duyck et al. (2004) mentioned a benchmark criterion where WordGen generated 80% pronounceable nonwords with appropriate settings. The *OT3* software was initialised using the BNC (*The British National Corpus*, 2001). In an exhaustive search, all five letter nonwords with an *OT3* value of more than 1.20 were generated. This resulted in 1073 nonword items ($OT3_{mean}$=1.40, $OT3_{max}$=2.56, $OT3_{min}$=1.20) where all items were legal, though about ten items appeared to be hard to pronounce (see Appendix C for a list of these items). This result challenges the benchmark of 80% pronounceable nonwords set by Duyck et al. (2004) using WordGen. Furthermore, *OT3* achieved this without referring to explicit phonological rules as Wuggy (Keuleers & Brysbaert, 2010).

Table 2.5: Comparison of *OT3* and word length and position specific mean log bigram frequency (*PLMLBF*). The fractional rank of *PLMLBF* in five letter words is displayed in FRank.

Nonword	*OT3*	*PLMLBF*	Word	*OT3*	*PLMLBF*	FRank
whrts	-2.13	3.19	*place*	1.47	3.19	0.08
tquld	-2.14	2.36	*alike*	1.68	2.36	0.77
qftch	-2.02	2.35	*nappy*	0.52	2.35	0.77
mythe	1.56	1.13	*rhyme*	0.11	1.11	0.99
atche	1.43	0.81	*hydra*	-0.57	0.84	1.00
answe	1.22	0.74	*pygmy*	-1.34	0.67	1.00

The results of generating nonwords using *OT3* and their respective typicality values were compared to the position and word length specific mean log bigram frequency (PLMLBF, results from N-Watch, Davis, 2005) in Table 2.5. Additionally, each nonword is shown with a word that corresponds to the respective *PLMLBF* value. For an orientation with regards to the *PLMLBF* values, all 3370 five letter words in the BNC (*The British National Corpus*, 2001) database were analysed using N-Watch and ordered by their *PLMLBF* values. The rank of each word was divided by the total number of five words resulting in the fractional rank which is displayed alongside in Table 2.5. These values showed that the *PLMLBF* values are not distributed equally across the spectrum, but the majority of items scoring in a range between above 2 and below 3. The *OT3* value of the first three nonword items indicated that these items are atypical in English and most likely unpronounceable. In contrast, the *PLMLBF* indicated that these items were comparable to normal English words, even though *nappy* and *alike* were amongst the more atypical words. Nevertheless, the examples show that *OT3* was not misled by some high frequency letters in specific positions and thus, more reliable in rating these items as atypical. At the other extreme, *OT3* predicts that the final three examples are typical in English. In contrast, the *PLMLBF* metric rates these items amongst the least typical percent of words or below. It could be argued that these items are not as atypical as the *PLMLBF* metric suggested. There was also some agreement between the two metrics, specifically in atypical words. Interestingly, *OT3* rated *alike* much higher than *PLMLBF*, indicating some disagreement in typical words. In summary, the *OT3* is more likely to pick up on trespassing orthographic rules than

PLMLBF (see Table 2.5). The next section reports experimental evidence regarding the psychological validity of the *OT3* metric.

2.3.6 Summary

The *OT3* algorithm for analysing and generating word and nonword stimuli operates language independently on a given corpus with word frequencies and an alphabet. The standardised measure provides a word length and language independent measure of typicality. Also, the resulting distribution of typicality results in a natural distribution that is found in RT data. This algorithm can also be used to analyse numerical code and symbols, e.g. a phonetic code. In comparison to WordGen (Duyck et al., 2004) the *OT3* metric is not based on the position specific or word length specific frequency of bigrams, but it uses the transition probabilities of trigrams. Furthermore, the *OT3* software provides an easy way of generating nonword items with a specified level of typicality. This contrasts with other nonword generator software, such as Wuggy (Keuleers & Brysbaert, 2010) where a base word is required for the process.

2.4 Experiment 1

This experiment aimed at testing the nonwords that were generated by the *OT3* algorithm. The nonwords were generated as highly typical and very atypical stimuli. Since the pronounceability of the nonwords forms a potential confound (Massaro et al., 1979; McClelland & Johnston, 1977), the experiment tested typicality within pronounceable and within unpronounceable nonwords in a standard lexical decision task. The word targets comprised high and low lexical frequency and typical and atypical word items covering a broad range. The usage of diverse word targets reflected the diversity amongst the nonword targets and was intended to hinder participants in relying on a legality judgement.

2.4.1 Methods

Participants. Twenty-eight participants from Royal Holloway, University of London took part in the study and either received credit points to complete their course requirements or £5 in exchange for their time. All were native speakers of English.

Table 2.6: *OT3* values of the nonwords used in Experiment 1. The values were derived by using the ECT frequencies in the CELEX database. Numbers in brackets indicate the items that included a trigram that was not present in the CELEX database.

		Pronounceability				
Typicality	Statistics	Pronounceable nonwords		Unpronounceable nonwords		Total
High	Mean	1.55	(0)	0.94	(0)	1.25
	Max	2.15		2.08		2.15
	Min	0.84		-0.11		-0.11
Low	Mean	-3.96	(11)	-3.88	(10)	-3.92
	Max	-2.54		-2.97		-2.54
	Min	-5.57		-5.14		-5.57

Stimuli & Design. 160 nonword targets were selected. Half of them had a high level of typicality and half had a low level of typicality. Additionally, half the items in each typicality class were pronounceable (e.g., *sument*, *udnop*) and half unpronounceable (e.g., *safght*, *vlpej*). Due to a lower density of unpronounceable highly typical items, the unpronounceable nonwords scored slightly lower in *OT3* than pronounceable items [$t(78) = 7.387$, $p<0.001$, $d=1.65$][2]. The low typicality items were matched across pronounceability [$t(57) = 0.478$, $p=0.634$]. The mean values are presented in Table 2.6. Each category comprised equal numbers of 5 and 6 letter stimuli. Importantly, none of the nonword stimuli had any orthographic neighbours according to N-Watch (Davis, 2005). A list with all items is provided in Appendix B.

The experiment also involved 160 word targets with half 5 and 6 letter words. The word targets were not matched on frequency (*mean*=148.1/million, *min*=0.3/million, *max*=3360.0/million; CELEX frequency according to N-Watch), but showed a considerable variance in frequency as well as in typicality ($OT3_{mean}$=0.33, $OT3_{max}$=2.39, $OT3_{min}$=-2.77; see Appendix B).

Procedure. All participants were tested in a quiet room in groups of up to three. They were handed the ethics consent form for the lexical decision experiment. Participants were told that words and nonwords would be displayed on the computer screen in front of them. They were instructed to indicate by pressing a clearly marked

2 effect size according to Cohen (1977, p. 44)

button of a button box whether the stimulus was a word or a not, responding as rapidly and as accurately as possible. The experiment started with ten practice trials, followed by 320 randomly ordered trials in four blocks separated by a self-paced break. The stimulus presentation and data collection was achieved by the use of DmDX (Forster & Forster, 2003). A Samsung CRT screen on a Windows XP PC was used.

For all presentation on screen a Courier font in size 24 was employed. Each trial started with a fixation asterisk presented at the centre of the screen for 500 ms. This was followed by the presentation of the target stimulus in uppercase. The target was presented until the participant responded or the response limit of 2000 ms passed. In case of failure to respond within the given time frame the participants saw a message indicating that their response was too slow. Feedback on the correctness of the response was provided during the practice session, but not with the experimental items.

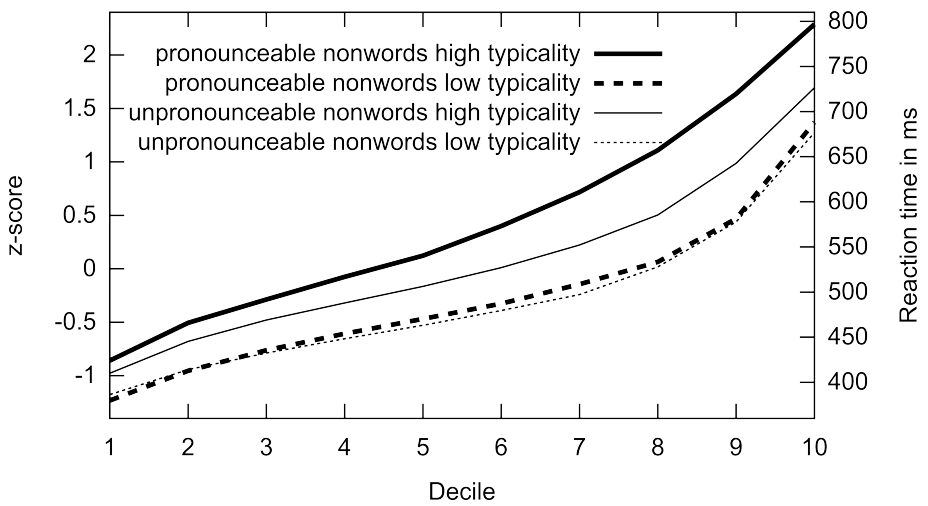

Figure 2.4: Plot of results of Experiment 1 by decile and condition. The scale of reaction times in ms is an approximation to reflect the actual corresponding value of z-score and reaction time.

Table 2.7: Mean reaction times in ms and error ratio in percent as a function of condition and typicality of Experiment 1.

Words	RT	514		
	Error ratio	4.55		
		Typicality		
		High	Low	*Effect*
Pronounceable nonwords	RT	581	498	*83*
	Error ratio	7.14	1.16	*5.98*
Unpronounceable nonwords	RT	535	490	*45*
	Error ratio	3.30	0.89	*2.41*

2.4.2 Results

Participants and items showing an error rate greater than 25% were dropped from the analysis as in all experiments in this thesis. This criterion did not affect participants, but seventeen words (*telex*, *epoch*, *idiom*, *xerox*, *idyll*, *typify*, *tawny*, *kiosk*, *topaz*, *codex*, *unripe*, *tract*, *cognac*, *vinyl*, *hyphen*, *lament*, *gauze*) and five pronounceable nonwords (*frount*, *govent*, *proust*, *prount*, *prould*) were excluded from the analysis. Unpronounceable nonwords were not affected by the exclusion criterion. Outliers were removed by 3 SDs for each participant using the correct responses. This affected 1.62% of the remaining data. The mean RTs and error ratios are shown in Table 2.7.

Repeated measures analysis of variance

ANOVAs were performed on the RT data of correct responses. Prior to the analyses the raw RT data were transformed to *z*-scores for each participant using the mean and standard deviation of all correct responses of the participant. This method was employed throughout the thesis. In all analyses the effect size according to Cohen's criterion (Cohen, 1988) was computed using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) by entering estimated η^2 in the direct computation. Cohen defined *f* values 0.1, 0.25, and 0.4 as small, medium, and large effects, respectively. The word targets were not of interest and thus, were not analysed further. As indicated in Table 2.6, the high typicality nonword conditions were not matched with regards to *OT3*. Thus, the effect of typicality was analysed within pronounceable and within unpronounceable nonwords separately. The effect of pronounceability was analysed

within low typicality nonwords and in a regression analysis accommodating the unmatched *OT3* values.

The separate analysis of nonword typicality in pronounceable nonword targets showed a significant main effect in *z*-scores [$F_1(1, 27) = 236.228$, $p<0.001$, $f=2.95$; $F_2(1, 73) = 117.403$, $p<0.001$, $f=1.27$], indicating that pronounceable nonwords with a low typicality were responded to faster than those with a high typicality. Similarly, the effect of typicality was significant in unpronounceable nonwords [$F_1(1, 27) = 95.020$, $p<0.001$, $f=1.88$; $F_2(1, 78) = 51.078$, $p<0.001$, $f=0.81$] where the low typicality items received faster responses than the high typicality items.

The analysis of pronounceability in low typicality nonwords revealed no significant effect of pronounceability [$F_1(1, 27) = 1.627$, $p=0.213$; $F_2(1, 78) = 1.000$, $p=0.320$]. A linear regression was performed using *OT3* as a predictor for the *z*-scores of RT across all nonword targets. The analysis revealed that *OT3* was a significant predictor for RT [$F(1, 132) = 143.305$, $p<0.001$, $r^2=0.521$]. Since the high typicality conditions were not directly comparable due to unmatched *OT3* values, the residuals of the regression were computed and are presented in Table 2.8. The residuals in the low typicality condition were very small and did not differ significantly between pronounceable and unpronounceable nonwords [$t(57) = 0.763$, $p=0.449$]. But there was a significant difference in the residuals of the high typicality conditions [$t(73) = 4.251$, $p<0.001$, $d=0.98$][3] indicating that the residuals were larger in pronounceable than in unpronounceable nonwords. That means after accommodating for the unmatched *OT3*

Table 2.8: Residuals of a linear regression using OT3 as predictor for *z*-scores of RT in nonwords in Experiment 1.

	Pronounceability	
Typicality	Pronounceable nonwords	Unpronounceable nonwords
High	0.15	-0.16
Low	0.03	0.00

3 effect size according to Cohen (1977, p. 44)

values, the unpronounceable nonwords received significantly faster responses than pronounceable nonwords.

Analysis in deciles

For a more thorough assessment of the typicality effect in the data, ten vincentiles for each condition were formed (Vincent, 1912). The fastest ten percent of the RT of one participant in each condition were assigned to the first bin, second fastest ten percent to the second bin and so on (Andrews & Heathcote, 2001; Balota et al., 2010; Balota, Yap, Cortese, & Watson, 2008; Ratcliff, 1979; Ratcliff & Murdock, 1976). In Figure 2.4 the mean of all participants per condition and bin is displayed, which reflects the behaviour of a single average (Ratcliff, 1979) or 'ideal' participant (Balota et al., 2008). The result of this procedure allows to test whether the typicality effect is present across the whole distribution of RT. A mixed model analysis was performed using R (R development core team, 2007) and lmer in lme4 (Bates, 2005; Bates & Sarkar, 2007) package (see Appendix E for a detailed list of results). In this analysis participants and items were entered as random factor and thus, the data was not aggregated prior to the analysis which increased the power of the analysis while maintaining the same α-error (Baayen, Davidson, & Bates, 2008). Specifically, the increased power was important because the vincentiling reduced the number of measurements per cell. Consequently, running separate analyses using either participants (F_1) or items (F_2) as random variable was obsolete. Two separate analyses for pronounceable and unpronounceable nonwords were computed. The *z*-scores of reaction time of the respective conditions formed the input for the analysis. The typicality effect within each decile was tested using a Markov chain Monte Carlo (MCMC) simulation. The output were *p*-values according the *t*-distribution and MCMC *p*-values (Baayen et al., 2008), where the latter tend to be more conservative (Brysbaert, 2007). The effect of nonword typicality was significant using either method. In all ten deciles in pronounceable as well as in unpronounceable nonwords typical nonwords received slower responses than atypical nonwords. This was the case in fast responses as well as in slow responses. The *t*-values increased in later deciles and the Ex-Gaussian analysis provided insight whether the effect size increased in slow responses. The analysis of pronounceability in low typicality nonwords revealed a

significant effect in decile seven and the slowest decile indicating faster responses to unpronounceable low typicality items. But the data also showed a tendency in fastest deciles for the opposite direction, i.e. faster responses to pronounceable low typicality nonwords. Thus, the analysis using the Ex-Gaussian fitting could show whether there is a systematic influence of pronounceability on τ.

Ex-Gaussian analysis

The distribution of reaction times is usually a positively skewed normal distribution (Heathcote, Brown, & Cousineau, 2004; Ratcliff, 1979; Ratcliff & Murdock, 1976). Following Ratcliff (1979; Ratcliff & Murdock, 1976) a convolution of a normal and an exponential distribution was chosen for fitting the raw data. There are various methods for deriving the according parameters of a distribution that describes the data (Cousineau, Brown, & Heathcote, 2004; van Zandt, 2000). The quantile maximum probability estimator (QMPE)[4] account (S. Brown & Heathcote, 2003; Cousineau et al., 2004) was chosen here. Following S. Brown and Heathcote (2003) the term *Ex-Gaussian* will be used to refer to the convolution of a normal and an exponential distribution. Using QMPE very accurate fits can be achieved (Rouder & Speckman, 2004) and effects that are masked in standard analysis may be revealed. Examples include effects in a semantic priming lexical decision task (Balota et al., 2008) and potential diagnostic uses in early stages of dementia (Alzheimer's type; Balota et al., 2010; Tse, Balota, Yap, Duchek, & McCabe, 2010). Also, Ratcliff and Murdock (1976) showed that the prediction of models can be falsified using the evidence from RT distributions, in particular when the mean RT does not provide evidence in favour of or against a particular model.

QMPE reveals three parameters per participant and condition describing the distribution of reaction times. The μ parameter reflects the central tendency of the Gaussian part of the distribution (the mean of the RT data were they normally distributed), σ denotes the standard deviation of the underlying Gaussian distribution,

4 Please note that the account was erroneously named *quantile maximum likelihood (QML)* by Brown and Heathcote (2003). Rouder and Speckman (2004) showed that the algorithm actually uses the *quantile maximum probability estimator (QMPE)*. The correct name was used in all later publications including the programme itself.

and τ is the slope of the distribution describing the positive skew. The algorithm splits the data of each participant in each condition in quantiles (or vincentiles) and the parameters μ, σ and τ are estimated for each quantile distribution. Finally, these results are used to estimate the distribution of RT for each participant and condition. After the estimation process the programme produces an exit code, where numbers smaller than 32 indicate accurate fits, 32 and 64 indicate some uncertainty and numbers greater than 64 indicate some data points might be missing (S. Brown, Cousineau, & Heathcote, 2008).

The QMPE software was used with the raw reaction time data after removing outliers. In analysing this and all other experiments the exit code was always smaller than 64, thus the estimates can be considered accurate (S. Brown et al., 2008) and exit codes were not reported. Mean values of the three parameters for each condition averaged across participants are shown in Table 2.9.

There was a significant main effect of typicality in pronounceable nonwords on μ [$F(1, 27) = 23.918$, $p<0.001$, $f=0.94$] reflecting faster responses in low typicality trials than high typicality trials. There was also a significant effect on τ [$F(1, 27) = 6.316$, $p=0.018$, $f=0.48$] reflecting a longer tail of the RT distribution in high typicality trials than in low typicality trials which was also supported by the visual inspection of Figure 2.4. This means that the difference between the conditions increased towards the slow tail of the distribution. There was no significant effect on σ [$F(1, 27) = 3.320$, $p=0.083$] where the tendency indicated a greater variance in high typicality trials. The analysis of unpronounceable nonword targets showed a significant effect on μ [$F(1, 27) = 20.972$, $p<0.001$, $f=0.88$] indicating faster responses to low typicality nonwords than to high

Table 2.9: Results of the Ex-Gaussian fitting using QMPE as a function of target pronounceability and typicality of Experiment 1.

	Pronounceable nonword			Unpronounceable nonword		
Parameter	High typicality	Low typicality	*Effect*	High typicality	Low typicality	*Effect*
μ	481	423	*58*	451	419	*32*
σ	53	39	*14*	39	30	*9*
τ	106	75	*31*	87	74	*13*

typicality nonwords. In contrast to pronounceable nonwords, there were no significant effects on σ [$F(1, 27) = 2.917$, $p=0.099$] and τ [$F(1, 27) = 2.280$, $p=0.143$]. The tendency in σ reflected a greater variance in high typicality trials than in low typicality trials.

The analysis of pronounceability in low typicality nonwords revealed no significant effects: μ [$F(1, 27) = 0.367$, $p=0.550$], σ [$F(1, 27) = 2.505$, $p=0.125$] and τ [$F(1, 27) = 0.037$, $p=0.850$].

2.4.3 Discussion

This experiment showed that nonword responses were significantly influenced by the typicality of the items. This was the case within pronounceable nonwords and within unpronounceable nonwords. Because there was a wordlikeness manipulation in word targets that mirrored that in nonword targets, the participants could not rely on wordlikeness for making their decisions. Rather, the experiment required participants to identify the words. That the word targets covered a broad range including difficult items was reflected in large number of errors to some word targets.

The manipulation of typicality showed a significant effect in unprimed lexical decision, even though the number of neighbours (*N*) was constantly zero in all nonword items. This contrasts with the use of summed bigram frequency where the effect was not stable if neighbourhood size was held constant (Andrews, 1992). The effect in the current experiment was also robust despite a pronounceability manipulation, because it was found in pronounceable and unpronounceable nonwords.

The effect of typicality was stronger in pronounceable nonwords, but the RT distribution in nonwords with a low typicality was very similar irrespective of their pronounceability (see Figure 2.4). The absence of a significant difference between pronounceable and unpronounceable low typicality nonwords could indicate that pronounceability did not show an effect per se, but rather added to the effect of typicality in forming very wordlike nonwords. The results were in contrast to earlier research where pronounceability showed an effect (Rubenstein et al., 1971). The

absence of an effect of pronounceability in low typicality nonwords was compatible with evidence observed with German stimuli (Loth, 2007). As in the current experiment, the lexical decision RT data showed a typicality effect in both types of nonwords. Low typicality nonwords received similarly fast responses independent of their pronounceability (pronounceable: 567 ms, unpronounceable: 570 ms). Furthermore, the effect of typicality was larger in pronounceable nonwords (567 ms and 731 ms) than in unpronounceable nonwords (570 ms and 634 ms). The current experiment and the previous findings in German (Loth, 2007) could indicate that pronounceability is not effective in low typicality nonwords, but pronounceability contributes to the wordlikeness in highly typical items.

The effect of typicality could be attributed to different levels of lexical activity that the nonwords in the different conditions triggered. Despite having no neighbours as measured in *N*, typical letter strings could have activated several word nodes that share letter combinations with the stimuli whereas this was not the case in atypical targets. A higher level of global lexical activity is more likely in word responses and thus, made it harder to reject the respective nonword targets. The effect size was very large with a difference of 80 ms in legal nonwords (f>1.00).

The more fine grained analysis of Experiment 1 revealed that the effect increased in slower responses. This was apparent in the increasing t-values in the analysis by deciles, but also in τ reflecting a significantly greater number of slower responses to highly typical compared to atypical pronounceable nonwords and a respective tendency in unpronounceable nonwords. An increasing effect size in slower responses is markedly different from the majority of the masked priming experiments that will reported in this thesis. In the current experiment, the targets differed between the conditions. Hence, the increase in effect size in slower responses could be attributed to a property of the targets themselves rather than to a change in the effect. In a network model, the rate of accumulating evidence in the decision channels would be different between nonwords in the same condition. This results in a different slope of increasing activity, which would be accompanied by an increase in variance. The results indicated a tendency for greater σ values in slower conditions, which supports this hypothesis.

Thus, the change in effect size across the RT distribution was not due to any processes that started at a later stage, but the speed of evidence accumulation which can be regarded as an intrinsic property of the targets.

The *OT3* measure of typicality was also used in the priming studies reported in the following chapters. This metric showed an effect in masked priming experiments and the current experiment where an effect in unprimed lexical decision in pronounceable and unpronounceable nonwords was demonstrated. These findings add to the reliability of the *OT3* metric of typicality.

2.5 Summary and Discussion

The review showed that the wordlikeness of word and nonword targets can influence the response times in lexical decision. The wordlikeness was approximated in various ways, including the number of neighbours *N* (Coltheart et al., 1977) and bigram frequency (e.g., Miller et al., 1954). Both, counting the number of similar words or using measures of orthographic redundancy resulted in facilitatory, null and inhibitory effects. Andrews (1992) argued that the different effects of neighbourhood density (*N*) in lexical decision could be explained by the same mechanisms that operate in different contexts. Grainger and Jacobs (1996) introduced a parameter in their computational model that weights global lexical activity and provided a theoretical account for facilitatory and inhibitory outcomes of manipulating neighbourhood density and bigram frequency. Grainger and Jacobs (1996) highlighted the importance of controlling for the wordlikeness of experimental items. Additional evidence for the effect of typicality stems from ERP (e.g., Hauk et al., 2006) and fMRI studies (e.g., Binder et al., 2003; Woollams et al., 2011). The experiments presented in this thesis also provided evidence that typicality was extracted from letter strings very quickly and influenced the state of the lexicon.

Even though typicality influences the presence, absence and direction of experimental effects (Andrews, 1989, 1992, 1997; Rumelhart & Siple, 1974), there is no conclusive measure of wordlikeness or typicality. C. J. Davis et al. (2009) showed that the standard definition of neighbourhood is too coarse for accommodating the effects

reported in their masked priming experiments. Other suggestions of measuring the neighbourhood of a word include the *OLD20* measure (Yarkoni et al., 2008). But increasing the complexity of neighbourhood measures reflects a move to a more global concept of wordlikeness, such as orthographic redundancy. Whilst neighbourhood reflects local connections and competition, a global wordlikeness measure reflects a comparison to the whole language and perhaps summed lexical activity. These effects can be isolated. In Experiment 1 and in the following experiments, it was argued that summed lexical activity triggered the typicality and prime-target congruency effects, whereas the inhibitory connections between neighbours were not relevant. In other experiments, e.g. in C. J. Davis and Lupker (2006), it could be argued that the inhibitory effects between two word nodes were very important.

The *OT3* metric for measuring orthographic typicality was suggested in this chapter. In contrast to other accounts, *OT3* provides a language independent measure of typicality expressed in standard deviations. Both the unit of measurement and the distribution of words is familiar to psycholinguists from *z*-scores and RT distributions (see Figure 2.3). Despite using no phonological information, the algorithm was able to produce 99% legal nonwords and challenges the benchmark set by WordGen (80%, Duyck et al., 2004). Furthermore, the *OT3* software allows generating nonwords in a specified band of orthographic typicality. In contrast to Wuggy (Keuleers & Brysbaert, 2010), *OT3* does not require a base word for nonword generation.

The unprimed lexical decision experiment supported the argument that typicality influences the recognition process independently of pronounceability. A strong effect of typicality was found in pronounceable and unpronounceable nonwords. Furthermore, there was no effect of pronounceability in low typicality nonwords. These findings provide support for the psychological validity of the suggested *OT3* algorithm.

2.6 Conclusion

The *OT3* algorithm provides a language-independent measure of typicality expressed in standard deviations. The effect of typicality as predicted by *OT3* was reported in eleven experiments of primed and unprimed lexical decision in two

languages. By providing this independent measure and resolving shortcomings of prior accounts, this account offers a reliable method of measuring typicality.

3. Experiment 2

3.1 Introduction

This experiment is aimed at investigating whether there is a response congruency effect in masked primed lexical decision that can be attributed to processes in the lexicon. The empirical evidence in the literature is unclear. Several studies have not shown such an effect (Norris & Kinoshita, 2008; Perea et al., 1998, 2010), but three studies reported evidence for congruency effects in lexical decision (Davis & Lupker, 2006; Jacobs et al., 1995; Klinger et al., 2000).

The review of theoretical accounts showed that there are several mechanisms that could result in a congruency effect. Some of these accounts attributed the congruency effects to processes other than the word recognition processes which are of particular interest in this thesis. For example, the results by Klinger et al. (2000) could be attributed to stimulus-response mappings due to repeated presentation of the stimuli. This bypasses the lexical processes of interest. In order to investigate processes in the lexicon rather than other mechanisms, it is necessary to design the experiment in a way that avoids other potential sources of apparent congruency effects. With regards to stimulus-response mappings, the targets in the current experiment were not repeated in order to avoid stimulus-response mappings. The formation of action triggers is another potential source of response congruency effects in lexical decision, but it has been identified as unlikely since the number of items is too large in order to prepare for specific items (Kunde et al., 2003). Similarly, lexical decision has been described as not automatic (Forster et al., 2003). Thus, the task is more likely to show effects of a lexical evaluation process rather than overlearnt response associations. Adhering to these restrictions will allow testing the impact of response congruency on lexical processes.

Prime-induced congruency effects are most likely in fast responses (e.g., Burle et al., 2002). In order to encourage participants to respond fast and increase the likelihood to find an effect, the stimuli were chosen so that the words and nonwords were relatively

easy to categorise. Targets were prototypical for their response category. According to the method outlined in Chapter 2.3 the words were chosen to have a high level of orthographic redundancy, whereas the nonwords scored low. The word targets were also high in frequency (e.g., *ROYAL*). The nonword targets (e.g., *AUDBC*) were required to have no orthographic neighbours (Coltheart et al., 1977). The primes were selected to be prototypical stimuli according to the same metrics (e.g., *order, dqrki*).

Under these conditions two accounts predict a response congruency effect. The deep processing account (Dehaene et al., 1998) suggests that the task instructions are applied to the prime unconsciously, which finally results in facilitation of congruent trials and interference in incongruent trials. The semantic overlap account (Quinn & Kinoshita, 2008) can predict a congruency effect, assuming that all meaningful letter strings, i.e. words, form a natural class and nonsense letter strings another. Finding an effect would support these two accounts, but challenge others such as the entry opening model (Forster, 1999). Furthermore, finding an effect would add evidence that congruency effects can emerge in lexical decision and command researching the empirical discrepancy outlined above.

3.2 Methods

Participants. Thirty-eight participants were recruited from the same population as in Experiment 1. All were native speakers of English.

Stimuli & Design. One hundred words and one hundred nonwords were selected for a masked priming lexical decision experiment to investigate a congruency effect between prime and target. The words were of a high frequency on the basis of CELEX English Corpus Type database (mean=336.5/million, min=24.47/million, max=2954.36/million; Baayen et al., 1995; N-Watch results) and were high in orthographic typicality as measured by the *OT3* algorithm ($OT3_{mean}$=1.30, $OT3_{max}$=1.99, $OT3_{min}$=0.58; see Chapter 2.3), e.g., *count*, *large*. The nonwords were selected to have no neighbours according to N-Watch (Davis, 2005) and to score low on the typicality measure ($OT3_{mean}$=-2.38, $OT3_{max}$=-0.81, $OT3_{min}$=-4.08), e.g., *dvnel*, *ftnzi*. All stimuli were formed of five letters. All stimuli are listed in Appendix B.

The words and nonwords were grouped in pairs of primes and targets. Every word and every nonword did double duty as target and prime, therefore the pairing procedure had to obey some prior set restrictions. In every prime-target pair none of the letters of the prime was allowed to occur in the target to prevent form priming effects. Each target was assigned a nonword prime and a word prime. Furthermore, it was made sure that a repetition of flipped pairs was avoided. In other words when a word primed a word (e.g., *large – YOUTH*) this relation was not simply flipped to form a pair where the former prime serves as target (giving: *youth – LARGE*), but another prime was used (e.g., *count – LARGE*).

In order to have every target appearing once per participant two versions of the experiment were designed. The words were ordered by frequency and assigned to the lists alternately with either the word or the nonword prime. The nonword targets also appeared in both lists, but their assignment was random. The ratio of incongruent and congruent trials was balanced in all lists to avoid strategic effects of prime evaluation due to the relatedness ratio (see Bodner & Masson, 2003; Bodner et al., 2006).

Procedure. All participants were tested in a quiet room in groups of up to three. They were handed written ethics consent forms and instructions for the lexical decision experiment. Participants were told that words and nonwords would be displayed on the computer screen in front of them. They were instructed to indicate by pressing a clearly marked button of a button box whether the stimulus was a word or a nonword, responding as rapidly and as accurately as possible. The experiment started with eight training trials, followed by 200 randomly ordered trials in four blocks separated by a self-paced break. The stimulus presentation and data collection was achieved by the use of DmDX (Forster & Forster, 2003). A Samsung CRT screen on a Windows XP PC was used.

For all presentations on screen a Courier font in size 25 was employed. Each trial started with a fixation asterisk presented at the centre of the screen for 500 ms. The anchor point for centring the presentation in this frame as well as in all other string presentations was the centre of the clear width of the pixel map occupied by the string. The fixation asterisk was followed by a blank screen for 200 ms. A row of five hash marks was presented at the centre of the screen for 500 ms serving as a forward mask to the prime. The prime was presented directly following the hash marks for three ticks

Table 3.1: Mean reaction times in ms and error ratio in percent as a function of lexicality and congruency of Experiment 2.

	Lexicality			
	Word		Nonword	
Congruency	RT	Error	RT	Error
Incongruent	498	3.70	520	3.80
Congruent	475	2.10	497	1.90
Effect	23	1.60	23	1.90

at 60 Hz. This is equivalent to a prime presentation duration of 50 ms, but depending on the actual screen device the refresh rate has some technical variance (e.g., 62 Hz resulting in 48.4 ms). The prime was presented in lower case. The font size of the prime was 75% of the target and the mask preventing ascenders and descenders from standing out from the forward mask and the target. Given the exact centring of the strings on screen masking of the prime was ensured. Directly following the prime, the target was presented in uppercase. The onset of the target also started the measurement of reaction time. Participants were required to respond within 3000 ms. Otherwise, they received feedback indicating that they responded too slow. During the training session participants received immediate feedback whether their response was correct. This helped the experimenter and the participants to ensure that task instructions were understood. The feedback was omitted in experimental trials to avoid interference from the presented words on screen.

3.3 Results

In all experiments reported here participants and items with error rates greater than 25% were dropped from the analysis. In Experiment 2, this resulted in the exclusion of one item (*visit*). In total 38 participants were tested, with 19 per list. Outliers were removed by 3 SDs for each participant using the correct responses. This affected 1.50% of the remaining data points. Results are presented in Table 3.1.

3.3.1 Repeated measures analysis of variance

An ANOVA was performed on the RT data of correct responses. Prior to the analyses the raw RT data were transformed to *z*-scores for each participant. As in Experiment 1 the *z*-scores were computed using all correct responses of a single participant and averaged per condition for the analyses. All analyses included the list factor as a random variable (Pollatsek & Well, 1995; Raaijmakers, Schrijnmakers, & Gremmen, 1999). There was a main effect of lexical status in *z*-scores [$F_1(1, 36) = 20.149$, $p<0.001$, $f=0.75$; $F_2(1, 195) = 57.494$, $p<0.001$, $f=0.54$] showing that words were responded to faster than nonwords. All further analyses of reaction times were performed for word and nonword targets separately.

A main effect of congruency was found in word targets in *z*-scores [$F_1(1, 36) = 53.331$, $p<0.001$ $f=1.22$; $F_2(1, 97) = 88.404$, $p<0.001$, $f=0.96$] showing that congruent trials received faster responses than incongruent trials. This main effect was also present in nonword targets [$F_1(1, 36) = 47.479$, $p<0.001$, $f=1.15$; $F_2(1, 98) = 72.192$, $p<0.001$, $f=0.86$]. The effect in nonwords indicated that congruent trials received faster responses than incongruent trials.

In error ratio, the main effect of response congruency was significant in the separate analyses of words [$F_1(1, 36) = 7.031$, $p=0.012$, $f=0.44$; $F_2(1, 97) = 8.968$, $p=0.003$, $f=0.30$] and nonwords [$F_1(1, 36) = 11.226$, $p=0.002$, $f=0.56$; $F_2(1, 98) = 14.907$, $p<0.001$, $f=0.39$].

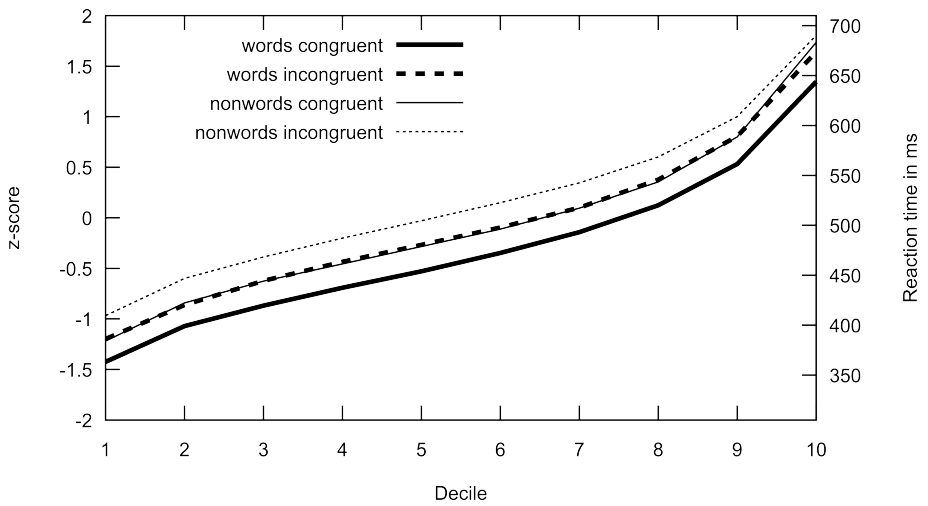

Figure 3.1: Plot of reaction times in Experiment 2 by deciles. The scale of reaction times is an approximation to reflect the actual corresponding value of z-score and reaction time.

In word and nonword targets, the congruent trials were responded more accurate than incongruent trials.

3.3.2 Analysis in deciles

Following the method outlined in Experiment 1, ten vincentiles for each condition were formed (Andrews & Heathcote, 2001; Balota et al., 2010, 2008; Vincent, 1912). In Figure 3.1 the mean of all participants per condition and bin is displayed, which reflects the behaviour of a single average participant (Balota et al., 2008; Ratcliff, 1979). A mixed model analysis was performed using R (R development core team, 2007) and lmer in lme4 (Bates, 2005; Bates & Sarkar, 2007) package (see Appendix E for a detailed list of results). In this analysis participants and items as well as list was entered as random factor. The input data were *z*-scores of reaction time from the separate analyses for word and nonword trials. The congruency effect within each decile was tested using a Markov chain Monte Carlo (MCMC) simulation following the method outlined in Experiment 1. The results indicated a significant advantage of congruent trials over incongruent trials in word and nonword targets in each decile. This indicated that in this experiment no shift of the effect occurred and the effect size was comparable throughout the whole RT distribution which was compatible with the visual inspection of Figure 3.1.

3.3.3 Analysis using the Ex-Gaussian fit

The distribution of reaction times was analysed using QMPE as outlined in Experiment 1. The input to the QMPE software was the reaction time data of correct responses after removing outliers. Mean values of the three parameters for each condition averaged across participants are shown in Table 3.2.

Table 3.2: Results of the Ex-Gaussian fitting of Experiment 2 using QMPE as a function of target lexicality and congruency.

	Word			Nonword		
Parameter	Incongruent	Congruent	*Effect*	Incongruent	Congruent	*Effect*
μ	430	414	*16*	457	431	*26*
σ	42	40	*2*	39	40	*-1*
τ	68	62	*6*	62	67	*-5*

The parameter estimates for each participant were analysed in a repeated measures ANOVA. A main effect of congruency was found in word targets on μ [$F(1, 36) = 11.608$, $p=0.002$, $f=0.57$] indicating faster responses to congruent than to incongruent trials. There was no effect on σ [$F(1, 36) = 0.228$, $p=0.636$] or τ [$F(1, 36) = 1.271$, $p=0.267$]. The analysis of nonword targets showed similar results. A congruency effect was present on μ [$F(1, 36) = 18.537$, $p<0.001$, $f=0.72$] showing faster responses in congruent compared to incongruent trials. There was no significant effect on σ [$F(1, 36) = 0.062$, $p=0.804$] or τ [$F(1, 36) = 0.951$, $p=0.336$].

3.4 Discussion

The experiment clearly showed a response congruency effect in word and nonword targets with an advantage of congruent trials over incongruent trials (see Table 3.1). According to the analysis using mixed models across deciles, the effect was present in each decile (Figure 3.1). This implied that the congruency effect was present in the whole distribution rather than affecting only parts of it, e.g. slow responses only (Balota et al., 2008). Using the Ex-Gaussian fitting algorithm (QMPE, Heathcote et al., 2004) the effect of congruency was found in the μ parameter. The other parameters of the Ex-Gaussian, the standard deviation of the normal distribution σ and the exponential part τ, did not reveal a significant effect of congruency. This means that the whole RT distribution was shifted: responses were faster in congruent trials than in incongruent trials. The congruency effect was also found in analysing the error score in word and nonword targets with fewer errors in congruent trials than in incongruent trials. All analyses of this experiment showed that there was a clear response congruency effect in both word and nonword targets. Furthermore, the effect size was large according to Cohen's *f* in word and nonword targets.

3.4.1 Empirical discrepancy

Previous empirical work was inconclusive on whether there is a response congruency effect in lexical decision. Some studies provided an indication that there was such an effect their data, even though the experiments were not explicitly aimed at finding a response congruency effect (Davis & Lupker, 2006; Jacobs et al., 1995). Studies that explicitly aimed at finding a response congruency effect did not report the

presence of such an effect in their data (Norris & Kinoshita, 2008; Perea et al., 1998, 2010) which was a critical piece evidence in favour of the Bayesian Reader model (Norris & Kinoshita, 2008). The current experiment clearly showed a response congruency effect in lexical decision. Potential reasons for this discrepancy include the experimental procedure, the difficulty of the task and differences in the primes.

Prime novelty and adaptation

One potential reason for the empirical discrepancy could be differences in the experimental procedure. The targets in this experiment were only presented once, but primes and targets were drawn from the same pool of items. Thus, about one half of the primes were *novel*, i.e. the participant had not encountered them as a target before, and the other half was *used* as prime after the participants responded to them consciously. The response congruency effect could have emerged as a result of learnt mappings between a stimulus and its response (Damian, 2001). In a post-hoc analysis each experimental trial was tagged as either *novel* or *used* with regards to the prime. Crucially, the stimulus-response mapping account predicts an interaction of prime novelty and congruency. That means a congruency effect is not expected in novel primes and strong congruency effects in used primes. The respective post-hoc analysis showed that the response congruency priming effects in novel and used primes were numerically almost identical and the interaction was not significant, although there was a main effect of prime novelty (see Table A.1, Table A.2 and Appendix A for more details).

In the course of the experiment, novel primes were more likely at the beginning whereas used primes were more likely towards the end of the session. The effect of adaptation or 'tuning into the task' could occur in the experiment and be mistaken as an effect of prime novelty as reported above. Most importantly, assuming that participants adapt during the experiment predicts that responses become faster after the initial trials. This contrasts with stimulus-response mappings predicting a null effect in novel primes. A post-hoc analysis was performed to test the hypothesis of adaptation.

The analysis of RT with regards to the four experimental blocks showed a main effect in nonword targets and a very strong tendency in word targets due to slower responses in the first block compared all other blocks (see Table A.3, Table A.4 and Appendix A for more details). The main effect of block indicated that participants adapted during the session, specifically after the first block. Due to the experimental procedure the largest ratio of novel primes was presented in the first block and the largest ratio of used primes in the last block. Thus, the main effect of block could have resulted in the main effect of prime novelty. Importantly, there was no indication that the congruency effect was absent or reduced in novel primes compared to used primes. Furthermore, the magnitude of congruency priming was similar in used and novel primes. Thus, the response congruency effect in this experiment was unlikely to be related to stimulus-response mappings. The presentation of primes as targets in one trial was unlikely to be the reason for the empirical discrepancy outlined above.

Difficulty of the task

The task difficulty is mainly determined by the choice of the targets. Grainger and Jacobs (1996) provided evidence suggesting that participants can adjust their decision criteria as a function of the nonword foils that are used in the experiment. In this experiment, the targets were easy to categorise for the participants and thus, a different set of decision criteria may have been used than in other experiments. Whether or not a response congruency effect only occurs if the stimuli are very easy to distinguish will be examined in Chapter 5. Also, the primes were very typical examples of their categories and could have provided more information than primes that are less clear candidates of a category, e.g. nonwords that are neighbours of words. The typicality of the primes will be addressed in a set of experiments as well (see Chapter 6).

3.4.2 Theoretical accounts

Deep processing account

The deep processing account (Dehaene et al., 1998) explains congruency effects by assuming the prime is processed in the same way as the target, though this processing

is incomplete. The deep processing account can accommodate the findings of the current experiment. An interesting prediction from this account is that the effect of prime congruency is independent from the difficulty of the explicit task. That means a congruency priming effect could occur in difficult targets as well as in easy targets. According to the deep processing account the task difficulty is not the underlying reason for the empirical discrepancy mentioned above (but see Grainger & Jacobs, 1996).

Semantic overlap model

The semantic overlap model (Quinn & Kinoshita, 2008) can explain response congruency effects in lexical decision as a result of semantic activation only if the monitor can form the summed activation of all semantic features. There was no systematic semantic relation between primes and targets in this experiment, but every word prime or target could activate semantic features which in turn could have contributed to the lexical decision. There is evidence that semantic properties of the targets, such as imageability (Cortese & Khanna, 2007; Morrison & Ellis, 2000), concreteness (James, 1975; but see Levy-Drori & Henik, 2006) and valence (Kousta, Vinson, & Vigliocco, 2009), can affect lexical decision latencies. Assuming that the total activity in semantic features can be monitored, a semantic account could accommodate the current findings.

3.5 Conclusion

The experimental data showed that a response congruency effect can emerge in masked primed lexical decision. This effect occurred with novel primes but also with primes that the participants had once encountered as a target. The magnitude of the priming effect was similar for word and nonword targets, but also for novel and used primes. The results challenge the assumption that congruency priming cannot occur in lexical decision (Forster, 2004).

The results also raised the question why this effect was absent in some studies (Norris & Kinoshita, 2008; Perea et al., 1998, 2010), but very clearly in present in the data of the current experiment. Reviewing theoretical accounts suggested reasons

including the target difficulty (cf. Grainger & Jacobs, 1996) and the informativeness of the primes. Both will be addressed in the following experiments.

4. Computational Models

In this chapter, I review computational models of word recognition and test these models using the data collected in Experiment 2. I discuss the Bayesian Reader (Norris, 2006; Norris & Kinoshita, 2008) and the Spatial Coding Model (Davis, 2010). I also discuss an adapted version of the latter model.

4.1 Bayesian Reader

The Bayesian Reader (Norris, 2006; Norris & Kinoshita, 2008) is a computational model of lexical access based on Bayes' theorem (Bayes & Price, 1763). The model can provide predictions for lexical decision (Norris, 2006) and masked primed lexical decision experiments (Norris & Kinoshita, 2008), but also the same-different task. In a lexical decision task the model evaluates the probability that the stimulus is a word and performs a lexical decision when this probability hits a threshold. The evaluation of the probability is based on accumulated noisy evidence at each time step. The model predicts the mean reaction time and error ratio per item.

4.1.1 Overview

Lexical decision in the model is not based on the identification of a single word, but on the overall evidence that the stimulus is word or a nonword. This process involves a comparison of two conditional probabilities: the probability that the stimulus is a word $P(X|is\ a\ word)$ or P_{word} versus the probability that it is a nonword $P(X|is\ a\ nonword)$, where X reflects the evidence from the input. As soon as one of these probabilities hits a threshold a response is triggered.

At the beginning of each trial the model is reset and the word priors reflect the respective word frequency. Each letter of the input is represented by a binary, position specific 26 element-vector. Hence, all representations of a letter string denote a point in a length-by-26-dimensional perceptual space. Due to this mechanism, the model can use a lexicon with words of a specific length only. At each cycle, the model receives a sample of the stimulus, which reflects the stimulus plus Gaussian zero-centred noise. The mean location and standard error of the mean are updated with each new sample.

These samples are used to compute a probability density function for each word, specifically the probability that the input was generated by a particular word $P(I|W)$; this probability is referred to as a likelihood. Following Bayes' rule, the posterior probability of the word – denoted $P(W|I)$ – is computed by multiplying the likelihood $P(I|W)$ by its prior probability $P(W)$ and dividing by the sum of all priors multiplied with their respective likelihood. This procedure results in the identification of a word that is close to the presented input, because $P(W|I)$ reflects a relative probability. Additionally, a transformation of all word probabilities and a probability for the stimulus being a nonword is required in order to compute the probability that the stimulus is a word (denoted by P_{word}).

To make a lexical decision, the model computes the probability that the stimulus is a word and the probability that it is a nonword. P_{word} is found simply by summing $P(W|I)$ across all of the words in the model's lexicon. The computation of the probability that the stimulus is a nonword is more complex, and depends on the assumption of a virtual nonword and background nonwords. Specifically, the virtual nonword is positioned close to the input in perceptual space, but never closer to a word than a given nonword distance. The probability that the input matches one of those nonwords is computed in the same way as for words, since the equation is blind to the lexical status of letter strings, but also to whether the string is pronounceable or legal. Eventually, the probability that the stimulus is among the background nonwords is approximated as $1P_{word}$.

The processing for prime and targets does not differ. At target onset, the basis for drawing samples is exchanged for the target and the samples drawn from the prime are deleted, i.e. the position in perceptual space extracted from the samples is defined by the target only. Priming results in a change in the priors, which are updated continuously during the processing of prime and target. If the prime and the target are orthographically related (e.g., *brown – CROWN*), the prime will have increased the priors of similar words (including *CROWN*). Hence, a priming effect is predicted. If the prime and the target are related by their response only (e.g., *quiet – CROWN*), the prime will not have altered the priors in the perceptual space around the target. Thus,

when the stimulus changes from the prime to the target, the influence of the priors around the prime in perceptual space is wiped out quickly, due to the multiplication with near-zero likelihoods (for these primed words, the likelihood $P(I|W)$ will be approximately zero, as they share no letters with the current input). Consequently, no response congruency effect is predicted.

Norris and Kinoshita (2008) showed that simulations of the Bayesian Reader correctly predicted no prime congruency effect for the stimuli tested in their experiment. They argued that the reason for this prediction is that the prime increases the evidence for words in a broad area of perceptual space, but does not alter the probability of a word or nonword response. However, it seemed plausible that the more informative primes used in Experiment 2 might result in large differences in P_{word} for word versus nonword primes. This possibility was tested in the following simulation.

4.1.2 Simulation

The stimuli used in the simulation were the same as in Experiment 2. Parameters were set to default level (see Appendix D). The decision threshold for a word decision was set to 0.9 and for a nonword decision to 0.01, as in Norris and Kinoshita (2008). The results of the simulation showed that the reaction times of the model did not show a congruency effect for either word or nonword targets, in contrast to the empirical data (see Table 4.1).

Table 4.1: Reaction time of the Bayesian Reader in cycles and empirical data of Experiment 2 in ms as a function of priming condition and target lexicality. The table includes results from identity and one letter-different control simulations.

	Bayesian Reader (simulation)		Empirical data (Experiment 2)	
	Target lexicality		Target lexicality	
Condition	Word	Nonword	Word	Nonword
Identity	*342*	*576*		
One letter-different	*348*	*573*		
Incongruent	378	558	498	520
Congruent	379	562	475	497
Effect	-1	-4	23	23

In order to achieve the one letter-different condition, the middle letter of the five letter stimuli was exchanged to form an illegal nonword. Note that each item was averaged across 50 simulations before contributing to the means (Norris, 2006).

As noted above, it seemed plausible that the relatively informative primes used in this simulation would produce large differences in P_{word} for word versus nonword primes. In fact, this prediction was correct as can be seen in Figure 4.1. The lefthand side of the graph (times less than zero) reflects the time when the prime is being processed. The model quickly distinguished between word (*word congruent*, *nonword incongruent*) and nonword primes (*word incongruent*, *nonword congruent*). Nevertheless, this did not result in a congruency effect, because the early difference in P_{word} is wiped out with the onset of target processing. The righthand side of the graph (times greater than zero) in Figure 4.1 reflects the time when the target is being processed. The values of P_{word} in Figure 4.1 differ by lexicality but not by congruency. The Bayesian formulation requires that priors are multiplied by likelihoods for each hypothesis (i.e., for each word in the model's lexicon). At the moment of switching from the prime to target, the target's likelihood is approximately zero and hence any priming effect preserved in the prior is wiped out by a multiplication with the likelihood. Also, there is no trace of the prime conserved in the system. From target presentation onwards the likelihood of the prime is approximately zero resulting in a prior of approximately zero in the next cycle. Thus, the prime's contribution to the pooled probabilities is wiped out as well. Hence, there is no trace of the prime after

target onset. Even if this process was more smooth, the model computes P_{word} from the current state of the lexicon at each time step, but the probability for the stimulus being the prime is approximately zero after a few hundred samples drawn from an orthographically unrelated (i.e., primes and targets that share no letters) target. It follows that the model cannot predict any priming effects (positive or negative) for primes that are orthographically unrelated to the target. This prediction directly follows from the model's fundamental assumption about how information is integrated during masked priming. This is consistent with the data from the mean values in Table 4.1. Thus, empirical data demonstrating a congruency effect for unrelated primes (Experiment 1; Jacobs et al., 1995) challenge the basic assumptions of the Bayesian Reader model.

In contrast, the model can predict priming in orthographically related primes and targets. In order to check this, two control conditions were created by pairing the target stimuli of Experiment 2 with new primes. First, an identity priming condition was simulated; the empirical data show strong priming effects when word targets (but not

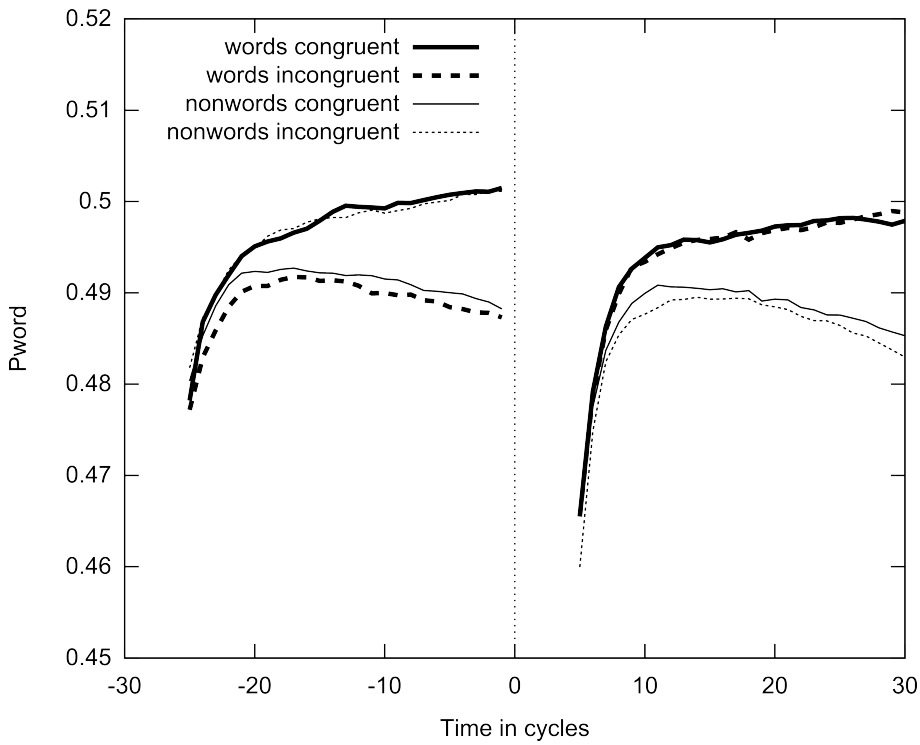

Figure 4.1: Bayesian Reader simulation of P_{word} using the items from Experiment 1. The negative numbers reflect cycles with the prime visible and positive numbers with the target visible. Each item contributed the average of 50 simulations to the means for each data point as recommended by Norris (2006), nevertheless the lines do not perfectly overlap due to the stochastic element in the model. The model does not output values for P_{word} in the first five cycles of prime and target processing which results in the discontinuity in the lines.

nonword targets) were preceded by identity primes (Forster & Davis, 1984; Forster et al., 2003). Secondly, a one-letter different condition was simulated where the primes were created by replacing the middle letter of the five-letter stimuli; such form primes also reliably produce priming in word targets (e.g., Davis & Lupker, 2006; Forster et al., 1987). The results of these simulations can be found in Table 4.1. The word targets showed large priming effects for both identity and one-letter different primes, relative to unrelated primes. However, there was no priming for nonword targets; the numerical difference between one-letter different and identity primes in nonword targets can be attributed to noise.

4.1.3 Conclusion

The Bayesian Reader model does not predict a response congruency effect. This prediction of the model was previously demonstrated by Norris and Kinoshita (2008). However, the reason for this prediction differs from the explanation given by Norris and Kinoshita (2008). It is not that the prime does not alter the probability of a word or nonword response. On the contrary, the computed probability of a word (P_{word}) was considerably greater following a word prime than following a nonword prime (see Figure 4.1). However, after target onset P_{word} is equivalent for the congruent and incongruent word trials, and likewise for congruent and incongruent the nonword trials. The reason is that priming is used to affect word priors, but these priors must be multiplied by likelihoods, which will be near zero for the prime words once the stimulus changes to an unrelated target. By contrast, in the case of related primes and targets the likelihood is not zero and priming can occur. This was shown by simulating an identity and a one-letter different control condition (Table 4.1). Although, the model's prediction of a null effect of congruency priming in words and nonwords is consistent with the authors' own empirical data (Norris & Kinoshita, 2008, Experiment 1) and data gathered in Spanish (Perea et al., 1998, 2010), it disagrees with the results of Jacobs et al. (1995) and Experiment 2. The latter data challenge a basic assumption about the integration of information in the Bayesian Reader.

4.2 Spatial Coding Model

The Spatial Coding Model (SCM; Davis, 2010) develops the Interactive Activation Model (IAM, McClelland & Rumelhart, 1981) that also provides the basis for the lexical component of other influential computational models, including the Multiple Read-Out Model (MROM, Grainger & Jacobs, 1996), the Dual Route Cascaded Model (DRC, Coltheart et al., 2001) and the Connectionist Dual-Process Model (CDP; Zorzi, Houghton, & Butterworth, 1998). In contrast to the IAM, the SCM uses a more flexible coding scheme for letter positions. This enabled the model to account for some shortcomings of a fixed slot coding scheme as recent work showed (Davis, 2010). Simulations of benchmark experiments showed that the SCM provides an excellent fit to empirical data, in particular to masked priming experiments (Davis, 2010). Conclusions about the lexical component of the model can generalise to other models that use a similar lexical component, and hence, I focus on the SCM here.

4.2.1 Overview

The SCM bases its lexical decisions on the opponent process model (Davis, 1999, 2010). According to that model, there are separate channels that accumulate evidence for *yes*- and *no*-responses respectively. These two channels are mutually inhibitory. A response is triggered when one of the channels hits a threshold. The information feeding into the *no*-channel is a constant, non-specific signal that builds up as the decision process continues over time. The *yes*-channel receives two inputs from the lexical level of the model. The first input represents the summed lexical activation, which is a measure of the global wordlikeness of a stimulus. The second input is information about specific word identification events. The relative contribution of these two sources is assumed to be subject to strategic factors. For example, when the nonwords in an experiment are relatively unwordlike (as was the case in the experiment simulated here) faster reaction times can be attained by placing a greater reliance on summed lexical activation than unique lexical identification (cf. Grainger & Jacobs, 1996).

At the start of each trial the model's activitions are reset. First the letters of the stimulus are recognised and the input is compared to the words in the lexicon. Each word is represented by a template that defines the letters in their position. By using the spatial coding scheme the input is compared to the word templates and a match value is computed. A match value is a measure of similarity between two given letter strings based on the spatial coding scheme. It ranges between 0.0 and 1.0. The more similar the two letter strings are the higher the match value is, with 1.0 marking equality (see Table 4.2). According to the respective match a word node receives activation from the stimulus presentation. As outlined above, the summed activation of all word nodes in the lexicon forms one source of input to the *yes*-channel. All word nodes form a competitive network and can inhibit each other. The word node with the strongest activity is able to suppress activation in other word nodes most efficiently. Identification of a word is said to occur once the activity of the corresponding word node hits a specific threshold. This activates an identification signal which forms the second source of information to the *yes*-channel. If the stimulus is not similar to any of the word nodes, the activation in each node is small and hence the summed lexical activation is low as well. In very wordlike stimuli the summed lexical activation is greater and so is the input to the *yes*-channel. Hence, it is possible to make a *yes*-decision prior to the identification of the stimulus. When simulating masked priming, it is assumed that the activity of the *yes*- and *no*-channels is reset at target onset,

Table 4.2: Examples for match values as calculated by the spatial coding model as a function of stimulus and template (C. J. Davis, 2010).

Stimulus	Template	Match value
table	TABLE	1.00
tablet	TABLE	0.86
stable	TABLE	0.86
trail	TRIAL	0.89
teach	BEACH	0.71
scale	STALE	0.86
smile	STALE	0.71
lager	REGAL	0.18

otherwise the initial decision bias due to the prime is difficult to overcome. Thus, the impact of the prime is limited to its effect on lexical activity (i.e., any “headstart” due to a prime has its locus at the lexical level rather than at the decision channels). As soon as one of the decision channels hits the threshold of 0.8 the response is triggered.

Lexical decisions based on summed activation can be performed faster, but are less reliable than decisions based on an identification of the stimulus (Grainger & Jacobs, 1996). Experiment 2 used very clearly distinguishable stimuli. The nonwords were very unwordlike (e.g., *miytd*) and the words high in frequency and very wordlike (e.g., *crown*). The parameters were identical to the published version of the SCM (Davis, 2010), except for the parameters that weight the inputs to the response channels in order to accommodate for the specific stimuli (see Appendix D). The global activity parameter was set to 0.5 ($y_{global}=0.5$), the unique lexical identification parameter was set to 0.0 ($y_{id}=0.0$), and the parameter weighting input to the *no*-channel was set to 0.22 ($n_{letter}=0.22$).

Because the current implementation of the SCM is a deterministic model and the main interest is in reaction time, the results reported here are restricted to reaction time and error rates are not considered.

4.2.2 Simulation

The stimuli in this simulation were the same as in Experiment 2 and in the previous simulation. The results of the simulation are shown in Table 4.3. Note that C. J. Davis (2010) set the parameters of the model so that priming effects predicted by the model in cycles are directly comparable to priming effects in empirical data in ms.

4.2.2.1 Nonword targets

Input stimuli that are unwordlike nonwords produce negligible lexical activation in the model. Thus, when the trial consisted of a nonword prime followed by a nonword target there was very little input to the *yes*-channel, and hence relatively rapid *no*-decisions. By contrast, when the prime was a high frequency word there was a much greater degree of lexical activation, resulting in a larger input to the *yes*-channel, which then inhibited the *no*-channel. Consequently, *no*-decisions were 13 cycles slower for nonword targets following word primes than for the same targets following nonword primes. The direction of the effect is compatible with the results of Experiment 2, but the model slightly underestimated the effect observed in nonword targets.

Table 4.3: Reaction times of the Spatial Coding Model in cycles and empirical data of Experiment 2 as a function of target lexicality and priming condition.

	SCM (simulation)		Empirical data (Experiment 2)	
	Target lexicality		Target lexicality	
Condition	Word	Nonword	Word	Nonword
Incongruent	170	166	498	520
Congruent	169	153	475	497
Effect	1	13	23	23

Note that the simulation was performed by using the summed lexical activation as the only input to the yes-channel.

4.2.2.2 Word targets

The situation for word targets is slightly more complex. In this case, inputs to the *yes*-channel facilitate faster correct decisions. Word primes accomplish this by increasing total lexical activation. However, another effect of word primes is to inhibit identification of the target word. This effect is a consequence of the lateral inhibition that drives the selection process in competitive network models like the SCM. For example, when the trial is *quiet – CROWN*, the prime causes the *QUIET* word node to become activated and to start sending lateral inhibitory signals to all other word nodes, including the *CROWN* word node. This lateral inhibition, which is not present when the prime is a nonword like *miytd*, persists for a little while after the onset of the target, and effectively slows down the identification of the target. As a consequence, the model actually predicts a negative priming effect of response congruency on the speed of target word identification. The combination of the facilitatory effect of word primes on summed lexical activation on the one hand and the inhibitory effect of word primes on target identification on the other hand produces a null effect of response congruency for word targets. Summed lexical activity as a function of time (in cycles) is

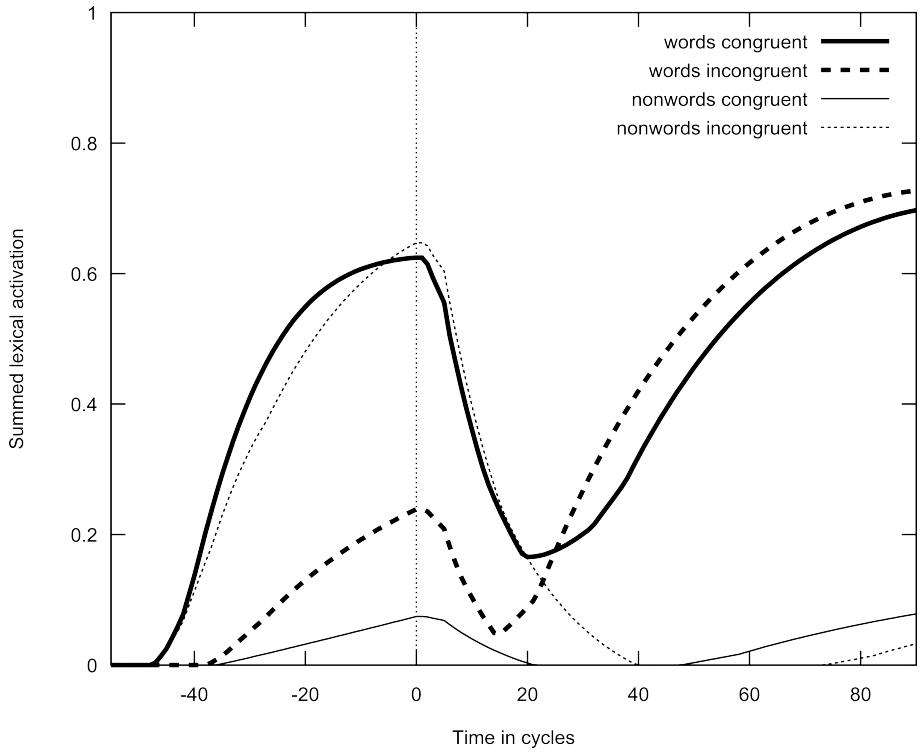

Figure 4.2: The flow of summed lexical activation in the Spatial Coding Model in word and nonword targets. The data was derived using the example quiet – CROWN (congruent) and miytd – CROWN (incongruent) for words and xvump - GIAGJ (congruent) and brown - GIAGJ (incongruent) for nonwords.

illustrated in Figure 4.2. Directly after target onset the activity in the congruent trial is greater than in the incongruent trial, but after a few cycles this pattern is turned around due to lateral inhibition between the word nodes corresponding to the prime and target words. The advantage in the early cycles of target processing and the disadvantage in the later cycles cancel out each other and result in the observed null effect. Note that the greater lexical activity associated with a congruent prime prior to target onset (the negative portion of the horizontal axis) does not contribute to the ultimate decision because of the reset of the *yes*- and *no*-channels that takes place when the target stimulus is presented (see Figure 4.7).

4.2.3 Inhibition in the Spatial Coding Model

The Spatial Coding Model (Davis, 2010) is unable to predict the response congruency effect reported in Experiment 2. Reviewing the simulations and the flow of summed lexical activation, the critical process was identified in the lexical component. Specifically, the interplay between word nodes during identification is critical. These processes are very similar to those in the Interactive Activation Model (McClelland & Rumelhart, 1981) as outlined above. The purpose of inhibition is to foster a process that narrows activation from a number of words down to a single word node. This aim can also be achieved by an alternative form of lateral inhibition called selective inhibition. I will explain this mechanism in more detail in the following section, and also present simulation results showing how this form of inhibition can account for response congruency effects. Furthermore, I will refer to a benchmark simulation highlighting the effectiveness of the modification.

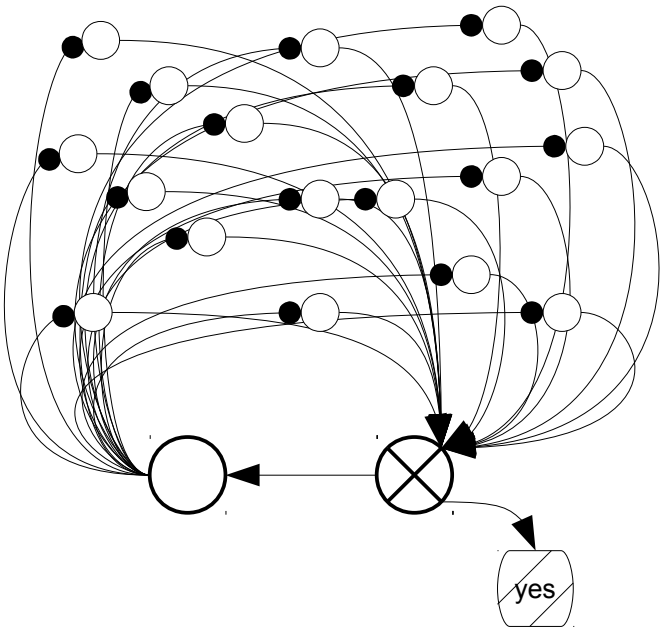

Figure 4.3: Homogeneous inhibition in an interactive activation model with a central unit to form the inhibitory signal to each word node.

4.2.3.1 Homogeneous inhibition

Simulations of the SCM reported in 4.2.2 followed the original IA model in assuming homogeneous inhibition between words of the same length. The SCM extended the IAM in implementing masking field principles. These principles increase the strength of inhibitory signals from nodes coding longer words. In homogeneous inhibition the word node for *CROWN* inhibits the node *QUIET* just as much as *CLOWN*. This is the case despite a strong similarity of *CROWN* and *CLOWN* and *CROWN* and *QUIET* not sharing a single letter. For computational convenience it was assumed that a single sum of lexical activity is formed. First, this sum serves as input for the *yes*-channel. Secondly, after some scaling this sum provides the inhibitory signal to all word nodes. This is depicted in Figure 4.3.

As soon as a word node becomes activated it sends an inhibitory signal to all other word nodes regardless of their relatedness. The inhibitory signals exchanged between prime and target in both directions are the reason for the absence of response congruency effects in the SCM (see Figure 4.4).

In conclusion, the SCM's ability to explain the data of Experiment 2 is hindered by inhibition between unrelated word nodes.

4.2.3.2 Selective inhibition

The homogeneous lateral inhibition that was assumed in the IAM and in the SCM is not the only option. Selective inhibition has been employed in the SOLAR model (Davis, 1999) and in a modified version of the IA model (Davis & Lupker, 2006). According to selective inhibition, only word nodes that code orthographically overlapping words send inhibitory signals to each other. Thus, the *QUIET* word node would not inhibit the *CROWN* word node, and hence word primes would exert no inhibitory effect on the identification of unrelated target words. The absence of this interference would allow a response congruency effect to emerge.

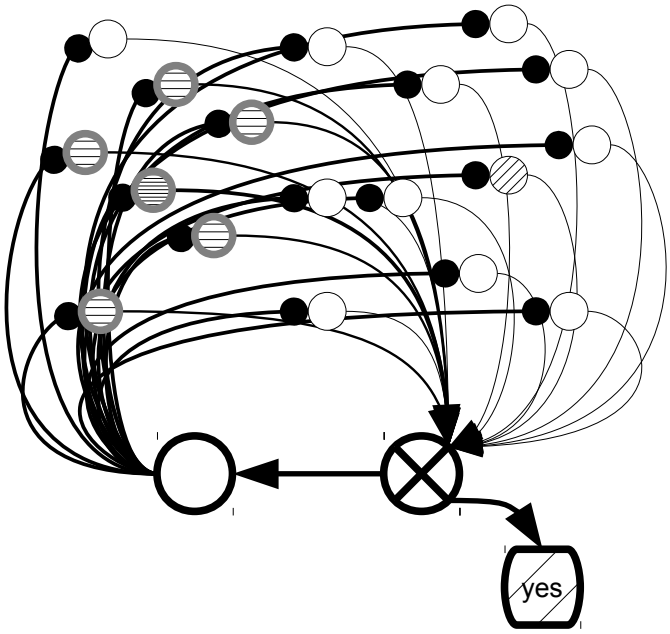

Figure 4.4: Prime activation in homogeneous inhibition. Emitting nodes and active connections are displayed stronger. Nodes in the prime cluster are hatched horizontally and the target node diagonally.

Overview

There are various different ways in which selective inhibition might be implemented. The two main problems are setting the weights of inhibitory connections and the computational expense of simulating specific connection strengths for each pair of word nodes. In the SOLAR model (Davis, 1999) it was assumed that connection strengths are a continuous function of orthographic similarity (e.g., that *CLOWN* inhibits *CROWN* more than *CHAIN* inhibits *CROWN*). A simpler approach was taken by Davis and Lupker (2006), who assumed binary inhibitory weights, and counted two words as orthographically overlapping if they shared at least one letter in the same position. For example, *AXLE* would receive inhibitory signals from word nodes like *ABLE*, *ARID* and *EXIT*, but not from word nodes like *DOOR* and *EMIT*. However, this position-specific approach is not consistent with the Spatial Coding Model, and leads to some irregularities when words of different length are considered. For example, *LATE* and *PLATE* are relatively similar according the SCM's coding scheme, even though they do not share any letters in the same (absolute) position. The following implementation of selective inhibition adopted aspects of both of the above approaches. The spatial coding scheme was used to determine the degree of similarity (Davis, 1999) and inhibitory connections were assigned binary weights (Davis & Lupker, 2006).

Connection weights

In the selective inhibition approach adopted here, two word nodes are either mutually inhibitory or independent, i.e. not linked by an inhibitory connection. Whenever the overlap between two words exceeded a criterion match value of 0.3 there was a bidirectional inhibitory connection between the corresponding word nodes. This implies that a word node like *CROWN* receives inhibitory connections from the *CLOWN*, *CHAIN*, *COW*, and *CLEAN* word nodes, but not from the *GROUP* word node. More importantly, there are no inhibitory connections between words that are orthographically unrelated, such as *QUIET* and *CROWN*.

Computational load

The number of inhibitory connections was limited in order to reduce the computational load. This was necessary because selective inhibition is notably computationally demanding, whereas homogeneous inhibition is less demanding; indeed, it was originally chosen for that reason (McClelland & Rumelhart, 1981). In a homogeneous inhibition account, the inhibitory signal to a word node can be determined by summing total word level activity. Additionally, self-excitatory input is added to a nodes activity in order to avoid self-inhibition. This kind of self-excitation is equivalent to subtracting the activity of each node from the total inhibitory signal. The computational advantage is that non-selective inhibition can be implemented on the basis of summed activity without modelling any lateral inhibitory interactions. In contrast, computing the inhibitory signal to a word node when selective inhibition is assumed necessitates multiplying each of the activities of the other word nodes by the inhibitory connections that these word nodes send to the node in question. That is, computing all of the lateral inhibitory signals at each step requires $O(n^2)$ calculations, where n is the number of word nodes. In the brain, such calculations can be performed in parallel, and so the value of n does not affect the speed of processing. When implementing selective inhibition on a serial computer, however, processing speed becomes highly dependent on the value of n. In relatively small networks the time cost is manageable, but in more realistically sized networks it becomes impractical to fully implement selective inhibition. For example, the network used in the simulations here encodes a lexicon of over 30,000 words, and thus would require a lateral inhibitory matrix of approximately one billion connections.

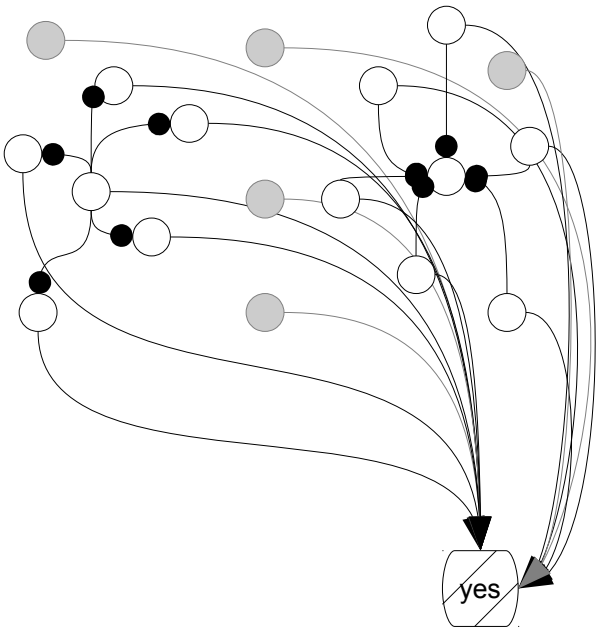

Figure 4.5: Network using selective inhibition. The nodes are the same as in Figure 4.3 and 4.4. The prime node on the left is displayed with its outgoing inhibitory connections, that do not reach the target node, that is shown with its incoming inhibitory connections. Nodes depicted in grey do not take part in the competition and are not simulated.

In practice, the vast majority of all word nodes are not involved in the selection process on a given input trial. Those word nodes do not receive any activation and hence, do not need to be exposed to inhibitory signals in order to ensure that only one word is identified. The word recognition process will not differ in its outcome if only those word nodes are taken into account that are similar to the input stimuli. This implies that a feasible way of solving the computational explosion caused by selective inhibition is to restrict the simulation on each trial to a subnetwork of critical nodes (i.e., those that will become active). This idea was implemented as follows. In order to select the critical word nodes on each trial, match values were computed for each word node in the lexicon relative to the prime and target stimuli. Hence, each word node in the network had two values assigned to it (e.g., its match with the prime *quiet* and its match with the target *CROWN*). For selection purposes, only the greater of the two values was taken into account, and the top 30 values were then used to select the critical word nodes. When the prime and the target are unrelated, as in Experiment 2 simulated here, this process results in two clusters around the location of each stimulus in lexical space. A network like this is presented in Figure 4.5. The inhibitory connections of the prime on the left do not reach the target node. Also, the figure

illustrates two clusters around the prime and target node. Nodes that are not taking part in the competition for the best match (i.e. are not part of either cluster) are not simulated. Computationally, this reduces the number of connections to an upper boundary of *900* and therefore the algorithm is in complexity class *O(1)*. This contrasts to $n*(n-1)$ many connections in $O(n^2)$ without forming a subnetwork.

4.2.4 Simulation in selective inhibition SCM

This simulation tested whether the modifications to the SCM were successful in enabling the model to accommodate the findings from Experiment 2. The homogeneous inhibition SCM already predicted a response congruency priming effect in nonword targets (Table 4.3). The results of the selective inhibition simulation demonstrated that the modification to inhibition in the lexical component enabled the model to correctly predict a congruency effect for both words and nonwords (see Table 4.4).

C. J. Davis (2010) demonstrates that the SCM is able to predict the priming effect in ms with a very high accuracy, but in this particular experiment the model underestimated the priming effect by approximately 10 cycles. Nevertheless, the model predicts a similar magnitude of priming in word and nonword targets which is compatible with the empirical data.

Table 4.4: Reaction time of the Selective inhibition SCM in cycles and empirical data of Experiment 2 in ms as a function of target lexicality and priming condition.

	Selective inhibition SCM (simulation)		Empirical data (Experiment 2)	
	Target lexicality		Target lexicality	
Condition	Word	Nonword	Word	Nonword
Incongruent	169	165	498	520
Congruent	156	152	475	497
Effect	13	13	23	23

Note that the simulation was performed by using the summed lexical activation as the only input to the yes-channel.

Two factors in a homogeneous inhibition account were identified that resulted in the prediction of a null effect in word targets. First, there was inhibition from the prime word node sent out to the target word node. This resulted in delayed activation of the target and a lower value of the summed lexical activation. Secondly, the word target node started to inhibit the prime cluster. This process resulted in a lower summed lexical activation as well. In a selective inhibition account the prime word node does not inhibit the target and vice versa. Thus, in a word trial the speed of activating the target is independent of the lexicality of the prime. Specifically, the negative impact of a word prime towards the target identification was removed. The same holds for inhibitory signals from the target node to the prime. Thus, the summed lexical activation is higher in a congruent word trial than in an incongruent word trial. This is illustrated in Figure 4.6. In congruent word trials, the summed lexical activation is consistently greater than in incongruent trials until about 30 cycles after target onset. Note, the decision channels are reset on target onset and the advantage of congruent trials is limited to the first 30 cycles of target processing.

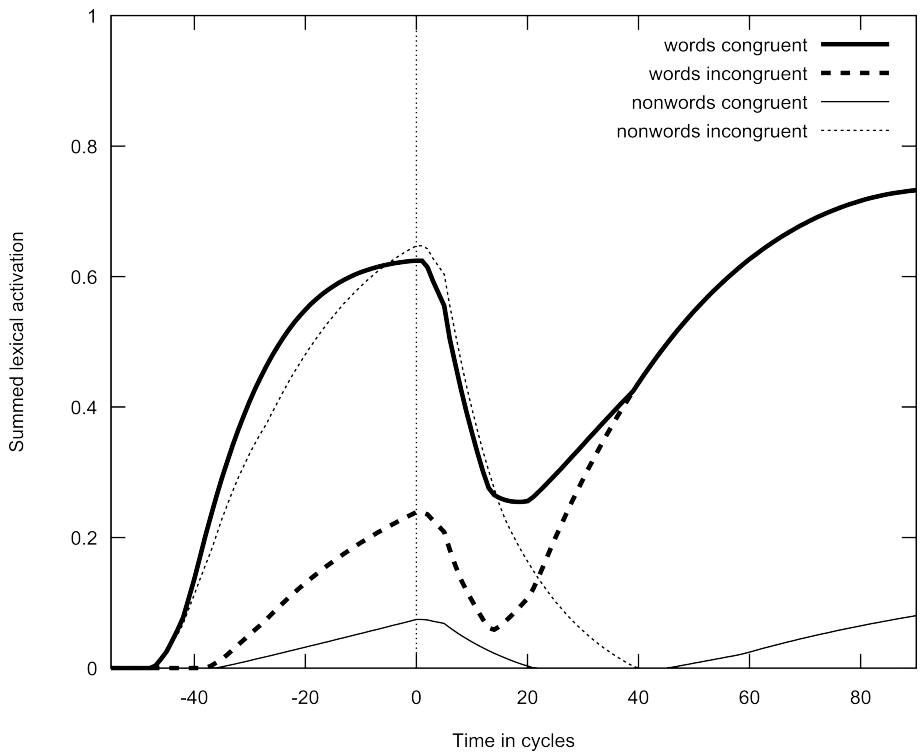

Figure 4.6: The flow of summed lexical activity in the selective inhibition SCM. The data was derived using the example quiet – CROWN (congruent) and miytd – CROWN (incongruent) for words and xvump - GIAGJ (congruent) and brown - GIAGJ (incongruent) for nonwords.

The effect of using selective inhibition in lexical component of SCM is also visible the *yes*-channel of opponent process model. The lefthand side of Figure 4.7 (negative cycle numbers) shows the the prime related activity. There is no difference between homogeneous and selective inhibition, which reflects that there was no difference at lexical level (compare Figure 4.2 and 4.6). As outlined above, the decision channels are reset with target onset. After target onset, the three lines reflecting a congruent and an incongruent word response using homogeneous inhibition and an incongruent response using selective inhibition are almost perfectly overlapping from about cycle 70 in Figure 4.7. Only the congruent trial using selective inhibition differs from the other three example decisions. This illustrates that additional evidence in favour of a word response was present in congruent trials, but only if selective inhibition was used.

4.2.5 Summary

The Spatial Coding Model (Davis, 2010) could not accommodate the response congruency priming effect reported in Experiment 2. The reason was found in the lexical component of the model, that is similar to the original Interactive Activation

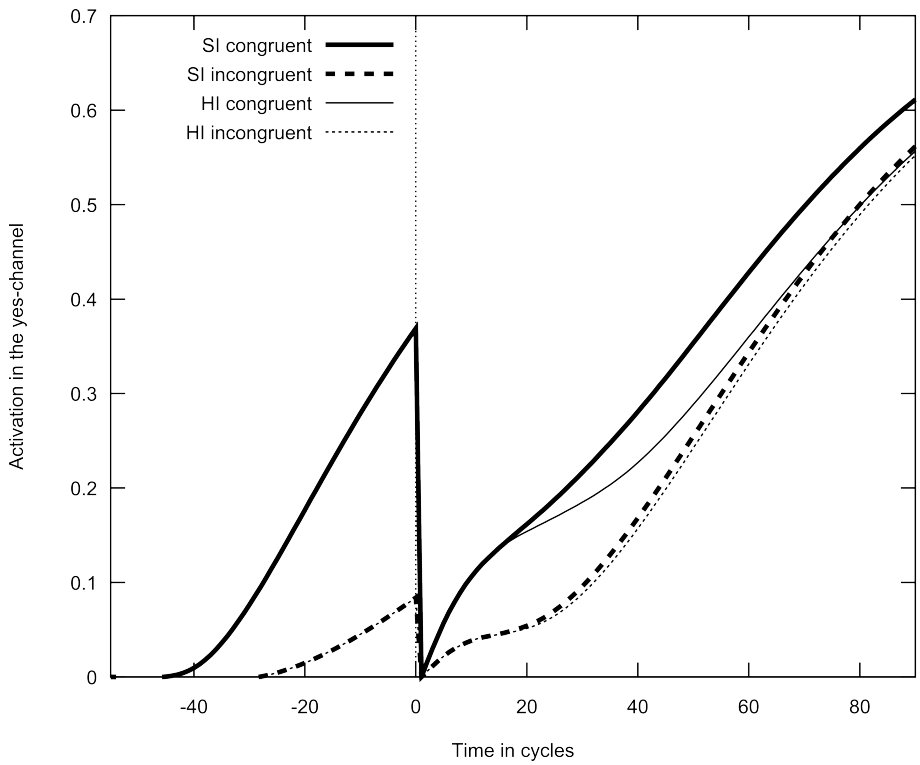

Figure 4.7: The flow of activity in the yes-channel of the opponent process model using homogeneous and selective inhibition. The data was derived using the example quiet – CROWN (congruent) and miytd – CROWN (incongruent) for words and xvump - GIAGJ (congruent) and brown - GIAGJ (incongruent) for nonwords.

Model (McClelland & Rumelhart, 1981). However, replacing homogeneous inhibition with selective inhibition enabled the SCM to predict the congruency effect reported in Experiment 2. Whether the modification affects the model's ability to accommodate other data is assessed in the next section. Also, the SCM with selective inhibition was used for a number of simulations in Chapter 8. In these simulations parameters of the model were fitted to the empirical data of Experiment 2 through to 10. Since the current simulation was aimed at highlighting the interactions between word nodes the unique lexical identification parameter was set to 0.0 ($y_{id}=0.0$), but the simulations in Chapter 8 include a number of experiments where the explicit lexical decision task was more difficult than in Experiment 2. Thus, the identification of a stimulus was set to default level and the results showed that the findings of Experiment 2 can still be captured by the model.

4.3 Benchmark simulation

The following benchmark simulations give an indication about the appropriateness of the models outside the context of response congruency. Most importantly, the selective inhibition account requires a test whether this assumption impaired the model's ability to predict other empirical data correctly.

Empirical evidence indicates that not all orthographic neighbours result in facilitatory effects on target recognition. C. J. Davis and Lupker (2006) showed that word neighbours used as primes result in inhibitory priming effects compared to unrelated word primes, whereas nonword neighbour primes result in facilitatory priming compared to unrelated nonword primes. The published version of the Spatial Coding Model predicts this outcome (Davis, 2010). However, the introduction of selective inhibition may have impaired the model's ability to capture inhibitory priming effects. This possibility was tested in the following simulation, which used the stimuli of C. J. Davis and Lupker (2006). The Bayesian Reader was also tested on the same stimuli.

Table 4.5: Prediction of the Bayesian Reader and Spatial Coding Models using the stimuli from Davis and Lupker (2006, Experiment 1) as a function of prime lexicality and prime relatedness in comparison to the empirical data.

Prime type	Word primes			Nonword primes		
Prime relatedness	Unrelated	Related	*Effect*	Unrelated	Related	*Effect*
Bayesian Reader	483	442	*41*	481	440	*41*
Homogeneous inhibition SCM	103	136	*-33*	104	86	*18*
Selective inhibition SCM	109	137	*-28*	107	86	*21*
Empirical data	609	633	*-24*	621	603	*18*

4.3.1 Bayesian Reader

The Bayesian Reader predicted facilitatory form priming effects as demonstrated above (4.1.2). However, the model did not distinguish between word and nonword primes (see Table 4.5). The underlying reason is that the target is similar to the prime and the likelihood of the target increased during prime presentation. As a result a facilitatory effect of priming emerges irrespective of the lexicality of the prime. This incorrect prediction confirms one of the criticisms of the Bayesian Reader model made by Bowers (2010).

4.3.2 Selective inhibition Spatial Coding Model

It was expected that the selective inhibitory mechanism would still be able to explain Davis and Lupker's (2006) findings. These priming effects depend on the interaction between word nodes that are similar to each other (e.g., *able* but not *door* inhibits *AXLE*). The absence of inhibition to dissimilar word nodes should not greatly influence the model’s ability to simulate inhibitory effects in neighbour priming. The results of the simulations (see Table 4.5) supported the hypothesis that the model provides a good fit to C. J. Davis and Lupker's (2006) data. This outcome suggested that selective inhibition increases the explanatory scope of the model without reducing its ability to explain other critical data.

4.4 Summary and Discussion

This chapter has reviewed two different approaches to modelling lexical decision data, and investigated the ability of those models to fit the data of Experiment 2. The Bayesian Reader (Norris, 2006; Norris & Kinoshita, 2008) cannot predict a response congruency effect due to the way information is integrated over time. The model failed to predict the facilitatory priming effect of prime-target response congruency, but was also unable to predict inhibitory priming effects (Davis & Lupker, 2006) in a benchmark simulation. The Spatial Coding model (SCM; Davis, 2010) did not accommodate the findings of Experiment 2, due to homogeneous inhibition in the lexical component of the model. After implementing selective inhibition the model accommodated the data and produced the correct prediction. Furthermore, the modified SCM still predicted inhibitory priming effects in benchmark simulation.

4.4.1 Homogeneous and selective inhibition

The SCM was modified in order to enable the model to predict a response congruency effect. Since the changes were not affecting the response channels which are still reset on target onset, the effect is not a matter of response congruency in the model. The move to selective inhibition allows lexical activity to be present in one part of the lexicon without affecting other, unrelated parts of the lexicon. Two separate clusters of activity produce a greater total lexical activity and this provides a stronger input to the *yes*-channel. The implementation of selective inhibition is economically reasonable, because only those nodes competing for final selection inhibit each other. This reduces the total amount of transmitted energy in a neuronal context, but it also reduces the computational load in a simulation. The SCM locates the congruency effect reported in Experiment 2 into the lexicon. This suggested some bias prior to the decision level. In fact, decision channels in the SCM are reset on target onset and the reason for predicting an effect is in the lexicon rather than in the decision component.

The experimental paradigm shed light on the interaction of word nodes that are unrelated to each other. This highlighted shortcomings of the homogeneous inhibition account and provided evidence in favour of selective inhibition.

4.4.2 Target difficulty

The targets used in Experiment 2 were particularly easy to categorise in words and nonwords. Specifically, the nonword targets were illegal letter strings and the word targets were comparably high in frequency. Thus, the experiment did not reflect a typical lexical decision experiment. It is possible that the response congruency effect that was observed in Experiment 2 could be attenuated in a more typical lexical decision experiment, but it could also be more pronounced. The increase in the difficulty level of the explicit task will reduce the response speed and accuracy (e.g., Dorfman & Glanzer, 1988; Usher & McClelland, 2001; Whaley, 1978).

Reduced effect

A smaller effect in a more typical lexical decision experiment could be attributed to two underlying processes. First, the information used in the lexical decision process can be sampled from different sources and their relative weights could shift. Secondly, the slower decision process could result in a reduced effect size.

Grainger and Jacobs (1996) suggested that in a harder lexical decision experiment a greater weight is assigned to lexical identification than to summed lexical activity. The latter was the source of the congruency priming in the simulation of Experiment 2 using the SCM (see 4.2). If summed lexical activity is assigned a lower weight, the

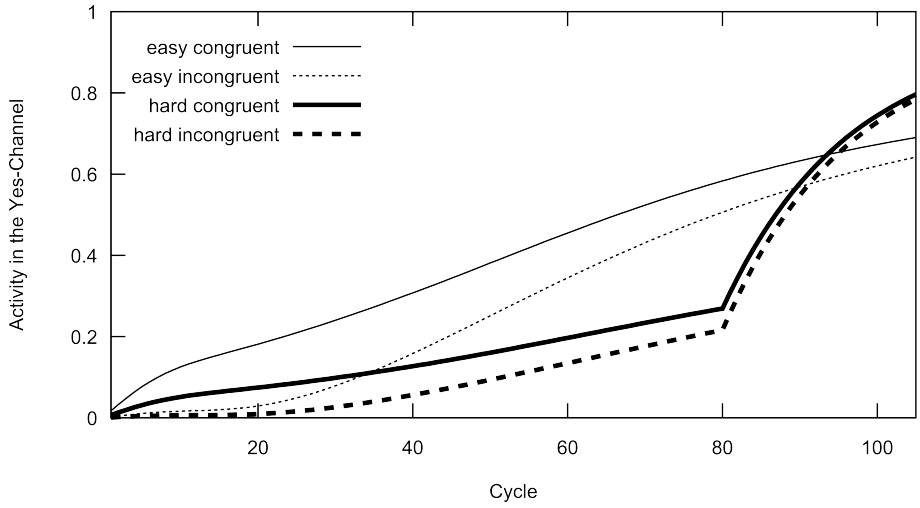

Figure 4.8: Comparison of the flow of activity in the yes-channel in the Spatial Coding Model in word targets. The data was derived using the example quiet – CROWN (congruent) and miytd – CROWN (incongruent).

difference between word primes and nonword primes in the *yes*-channel is reduced compared to the situation in Experiment 2. This would result in an attenuated congruency priming priming effect which would be hard to detect in empirical data. The example that was used for producing Figure 4.2 and 4.6 (*quiet/myitd – CROWN*) was simulated in different contexts. The activity in the *yes*-channel after target onset (i.e., after resetting the decision channels) is shown in Figure 4.8 with the specific adjustments for an easy ($y_{global}=0.5$ and $y_{id}=0.0$) and a hard lexical decision task ($y_{global}=0.2$ and $y_{id}=1.0$). The difference in the parameters reflects a shift from summed lexical activity to identification. Figure 4.8 shows that the critical difference between the congruent and incongruent condition are almost completely washed out as soon as the target was identified at around cycle 80.

Another potential reason for an attenuated effect is the duration of the decision process. In the simulations using the SCM, the difference between the word prime and the nonword prime condition was found in the early cycles of the target processing. In a more typical lexical decision experiment, the time difference between the occurrence of the effect and the actual decision would be increased. In the SCM (Davis, 2010), it was assumed that the decision is based on the average of the decision channel activity. Thus, in a slower decision process more cycles will feed into the average where the effect was not found and the response congruency effect would be attenuated.

Decision models

The decision process in two-alternative forced-choice tasks was modelled by decision models (Ratcliff, 1978; Smith & Ratcliff, 2004; Smith, Ratcliff, & Wolfgang, 2004; Zandt, Colonius, & Proctor, 2000). This group of models provides predictions about the reaction time distribution and error rates. Also, a prediction on whether the effect size in response congruency in lexical decision should increase, decrease or remain unchanged as a function of the target difficulty can be derived.

The leaky accumulator model (Usher & McClelland, 2001) shows similarities to the opponent process model that was used in the SCM (Davis, 1999, 2010). Both models use two integrators or response channels with one integrator for each response. These

integrators can mutually inhibit each other as function of their current activation level. In both models, a decision (or response) is triggered when one of the integrators exceeds a threshold. In contrast to the opponent process model, the leaky accumulator model shows leakage. That means the integrators are subject to spontaneous decay and, as a result, at each cycle a proportion of the activity is cancelled from the the respective integrator. This spontaneous decay is similar to the decay in the word nodes of SCM and thus, the leaky accumulator model provides interesting comparison to the SCM's decision model. In sum, activity in the integrators is influenced by three parameters: the incoming evidence, the leakage and the mutual inhibition between the integrators.

Two integrators were used by other decision models as well. The race model (Vickers, 1970) employed two integrators that are not connected through inhibition. Due to the absence of inhibition, the integrator that received a critical amount of evidence fastest triggers a decision. Shadlen and Newsome (2001; Ditterich, Mazurek, & Shadlen, 2003) introduced a decision model with feed-forward inhibition, i.e. incoming evidence is excitatory for one integrator and inhibitory for the other (counterevidence). Also, pooled inhibition that is similar to homogeneous inhibition at a lexical level was employed in decision models. Wang (2002) introduced such a model. Bogacz, Brown, Moehlis, Holmes, and Cohen (2006) showed that linearised versions of all these models are equivalent to the diffusion model (Ratcliff, 1978) under certain parametrisations. In contrast to the above models, the diffusion model uses only a single integrator. The diffusion model successfully simulated a number of empirical findings (see Bogacz et al., 2006 for review), including lexical decision (Ratcliff, Gomez, & McKoon, 2004).

As Bogacz et al. (2006) showed the scope of the decision models that were introduced above is comparable under certain parameters. For the purpose of the argument here and in Chapter 8, I refer to the leaky accumulator model. This model is similar to the opponent process model that is used by the SCM (Davis, 1999, 2010). The most important difference between the two models is the leakage.

Both hypotheses, a shift in weight assigned to summed lexical activity and slower decision process suggest that the presence of a response congruency effect could depend on the difficulty of the word-nonword discrimination. The analysis of the reaction time distribution that was applied in Experiment 2 could distinguish between the effect of these two processes. If the absence of a response congruency effect in a more typical experiment had to be attributed to a shift in the weights assigned to summed lexical activity and lexical identification, it is expected that the effect was attenuated across the whole distribution of RT. In case the effect was attenuated due to averaging or leakage, it is expected that effect was reduced in the slower responses, but present in the fast responses.

Increased effect

The effect size of the response congruency effect could also increase in a more typical experiment. In semantic priming studies size of effects attributed to interference increased in slow responses (Balota et al., 2008; Yap, Balota, Cortese, & Watson, 2006). Also, Bodner and Masson (1997) argued that a more effortful processing enables the prime to interfere stronger with the target. On grounds of these studies the congruency priming effect could increase with a harder task.

These accounts provide clear, but opposite predictions for an experiment with an increased task difficulty. If the response congruency effect was absent in a harder task the empirical discrepancy between Experiment 2 and other empirical data (e.g., Norris & Kinoshita, 2008; Perea et al., 1998, 2010) could be explained. The next chapter presents the experiments that tested these hypotheses. The respective empirical data was also simulated using the SCM with selective inhibition in Chapter 8 where the parameters denoting the influence of summed lexical activity and identifying a stimulus were fitted to the empirical data.

5. Effects of target difficulty

Experiment 2 clearly demonstrated a response congruency effect in lexical decision. This was compatible with one of the experiments in the literature (Jacobs et al., 1995), but it was in contrast to three other studies (Norris & Kinoshita, 2008; Perea et al., 1998, 2010). I have pointed out several factors in the introduction and in the discussion of the simulations that can contribute to the presence or absence of response congruency effects, including the difficulty of the explicit word-nonword discrimination task. The more difficult task could result in a shift of the weight that participants assign to summed lexical activity (Grainger & Jacobs, 1996) and by relying on identification the response congruency effect could diminish. The more difficult task could also result in a more effortful processing and thus, in an increased impact of the prime (Bodner & Masson, 1997). This chapter investigates the impact of the task difficulty on the response congruency effect.

5.1 Empirical discrepancy

Since the results are mixed across laboratories and languages, this may imply that the experiments differed in one or more variables. All experiments reported in the literature were masked primed lexical decision tasks. The task requirements were therefore very comparable. Thus, factors such as the category size appear not to be the critical variable. One important difference is the response speed in these experiments.

There appears to some relation between the mean RT of an experiment and the observation of a response congruency effect (see 1.4.2.3). Where the experiments with the faster RT tended to show a congruency effect (Experiment 1; Davis & Lupker, 2006; Jacobs et al., 1995), but experiments with slower responses did not (Norris & Kinoshita, 2008; Perea et al., 1998, 2010). The simulation of Experiment 2 in the Spatial Coding Model (Davis, 2010) showed that the advantage of congruent over incongruent trials is prevalent in the early cycles of target processing (see Figure 4.6). Since the model averages across cycles, the impact of these cycles is reduced with a longer processing time. A similar prediction could be derived from the leaky accumulator model (Usher & McClelland, 2001). Due to the leakage in the decision channels, the impact of the

prime is more likely to trickle out with a longer processing duration. Both models suggest that slower responses are associated with a smaller response congruency effect. This could explain the discrepancy in the empirical results. There is also empirical data from other experimental paradigms in support of the idea that priming effects are reduced in slower responses (Abrams, 2005; Burle et al., 2002; Greenwald et al., 2003; Kinoshita & Hunt, 2008). Following these predictions, slowing down the response speed of participants by making the explicit task more difficult could resolve the empirical discrepancy and highlight underlying principles. In contrast, there is also evidence for increasing priming effects with slower responses in semantic tasks (Balota et al., 2008; Bodner & Masson, 1997). The results can therefore provide evidence of what kind of theoretical framework is applicable and shed some light on the locus of the effect.

In order to slow down the response speed of participants the difficulty of the explicit task can be increased. Reviewing Norris and Kinoshita's (2008) English stimuli suggested that the items were harder to categorise than the stimuli used in Experiment 2. For example, the word targets were lower in frequency (9.2 per million) compared to Experiment 2 (336.5 per million). In addition, the nonword targets were more wordlike overall than those in Experiment 2. Also, the nonwords were similar to target words (*STEAK* vs *STERK*; *PLANT*, *PLANE* vs *PLART*). Other experiments in Spanish (Perea et al., 1998, 2010) and in French (Jacobs et al., 1995) are less straight forward to assess with respect to their stimuli. The aim of Experiment 2 was to use easy targets in order to maximise chances for finding an effect. The following experiments will make use of a harder explicit task.

In the following, I introduce three experiments where the task difficulty of the explicit lexical judgement was manipulated. In Experiment 3 the same primes and targets as in Experiment 2 were used, but the nonword targets were replaced by more wordlike items. Similarly, Experiment 4 used the primes and targets of Experiment 2, but the word targets were replaced by less frequent words. Finally, in Experiment 5 both word and nonword targets were replaced. The aim of the experiments was to find an explanation for the empirical discrepancy between strong response congruency

effects (Experiment 2) and null effects (Norris & Kinoshita, 2008; Perea et al., 1998, 2010). Furthermore, the prediction from computational models (Davis, 2010; Usher & McClelland, 2001) and empirical data (Burle et al., 2002) showing that the effect size decreases in slower responses, was tested.

5.2 Experiment 3

This experiment aimed to examine whether the prime congruency effect established in Experiment 2 is stable in tasks with an increased level of difficulty of the explicit lexical decision task and over all slower RTs. To achieve this, the nonword targets were replaced and made harder to reject compared to the ones used in Experiment 2.

5.2.1 Methods

Participants. Forty participants were taken from the same population as in Experiment 1.

Stimuli & Design. The same words were used as in Experiment 2. The nonword targets were selected to score higher in a typicality measure ($OT3_{mean}$=-0.53, $OT3_{max}$=0.51, $OT3_{min}$=-3.41; see Chapter 2.3). All of them were pronounceable and appeared to have at least a high potential to form an English word according to their orthotactics (e.g., *teirp*, *dulew*). As in Experiment 2, N-Watch (Davis, 2005) was used to ensure that the nonwords do not have neighbours. The nonword primes were selected to score very low on the typicality measure and were also checked using N-Watch (e.g., *qbnnj*, *zulmk*). All stimuli were formed of five letters and are listed in Appendix B.

As in Experiment 2, pairs of primes and targets were formed with the stimuli. In this experiment only the words functioned as both targets and primes, though the matching procedure was very similar. In every prime target pair, none of the letters of the prime were allowed to occur in the target. Each target was assigned a nonword prime and a word prime. The matching differed from Experiment 2 in so far that nonword primes and targets were taken from disjoint sets; therefore nonwords could not be involved in repeated flipped pairs. Again, every target could occur in a congruent and an incongruent condition. For counterbalancing purposes two lists were formed, where each target appeared once per list.

Procedure. The procedure was the same as in Experiment 2.

Table 5.1: Mean reaction times in ms and error ratio in percent as a function of lexicality and congruency of Experiment 3.

	Lexicality			
	Word		Nonword	
Congruency	RT	Error	RT	Error
Incongruent	514	4.30	559	2.75
Congruent	496	1.80	548	2.35
Effect	18	2.50	11	0.40

5.2.2 Results

Participants and items with an error rate greater than 25% were dropped from the analysis. This criterion affected none of the participants or items. Half of the 40 participants were assigned to list 1 and list 2. Outliers were removed by 3 SDs for each participant using the correct responses. This affected 1.50% of the data points. The mean reaction times in ms and the error ratio are shown in Table 5.1.

Repeated measures analysis of variance

An ANOVA was performed using the participant-wise *z*-scores of correct responses. The list factor was always added as between participants. Congruency entered the analysis as a repeated factor within participants and lexicality was a repeated within participants factor in F_1 and a between items factor in F_2. The analysis revealed a main effect in *z*-scores [$F_1(1, 38) = 109.234$, $p<0.001$, $f=1.70$; $F_2(1, 196) = 53.658$, $p<0.001$, $f=0.97$]. This indicated that words were responded to faster than nonwords. All further analyses were performed for word and nonword targets individually.

In word targets, a significant main effect of congruency was observed in *z*-scores [$F_1(1, 38) = 22.584$, $p<0.001$, $f=0.77$; $F_2(1, 98) = 32.848$, $p<0.001$, $f=0.58$] indicating faster responses to congruent than to incongruent word trials. This was also the case in nonwords in *z*-scores [$F_1(1, 38) = 9.781$, $p=0.003$; $F_2(1, 98) = 20.937$, $p<0.001$, $f=0.46$] where congruent trials received faster responses than incongruent trials.

The analysis of error data revealed a significant main effect of congruency in word targets [$F_1(1, 38) = 17.541$, $p<0.001$, $f=0.68$; $F_2(1, 98) = 18.754$, $p<0.001$, $f=0.44$] reflecting more erroneous responses in the incongruent than in the congruent condition. There was no effect in nonword targets [$F_1(1, 38) = 0.628$, $p=0.433$; $F_2(1, 98) = 0.672$, $p=0.414$].

Analysis in deciles

As in Experiment 1 the reaction times were binned in ten vincentiles and the results submitted to an analysis using lme4 in R (see Appendix E for a detailed list of results). The results showed that the congruency effect was significant in each decile in word targets. This suggested a stable advantage of congruent over incongruent trials across the whole RT distribution, which is also evident from the graph in Figure 5.1. In nonword targets there was no effect of congruency in decile eight, but there was a significant effect in all other deciles. The effect in the deciles one to seven was positive. That means congruent trials elicited faster responses than incongruent trials. Interestingly, slower responses that were binned in deciles nine and ten showed a negative effect. This means that incongruent trials received faster responses than congruent trials. Figure 5.1 illustrates this pattern. The lines for congruent and incongruent nonword trials cross at decile eight where no significant effect was found. In total, 70% (seven deciles) of the nonword trials showed a positive effect, 20% a negative effect and 10% no effect at all. This pattern still resulted in a significant positive effect in mean reaction times.

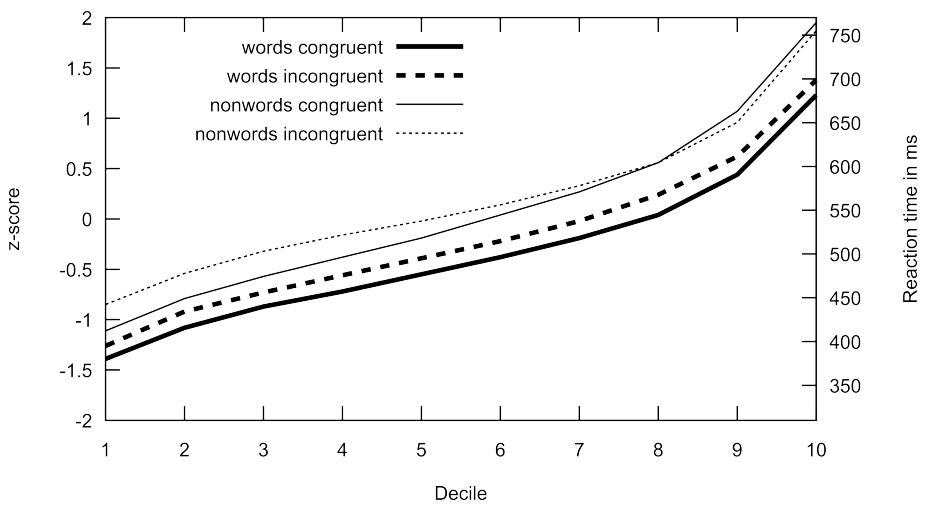

Figure 5.1: Plot of the results of Experiment 3 by decile and condition. Note that the scale in ms is an approximation and data was plotted by z-scores.

Ex-Gaussian analysis

The fitting of the Ex-Gaussian parameters was performed according to the method outlined in Experiment 1. The results are shown in Table 5.2. There were no significant effects in word targets: μ [$F(1, 38) = 2.678$, $p=0.110$], σ [$F(1, 38) = 0.293$, $p=0.592$] and τ [$F(1, 38) = 0.942$, $p=0.338$]. The analysis in nonword targets showed a significant effect of congruency on two parameters: μ [$F(1, 38) = 12.254$, $p=0.001$, $f=0.57$] reflecting faster responses in congruent trials, and τ [$F(1, 38) = 5.199$, $p=0.028$, $f=0.37$] indicating a longer tail of the RT distribution in congruent than in incongruent trials. This is depicted in Figure 5.1, where the lines in the decile plot cross due to the smaller number of very slow responses in the incongruent condition. There was no effect in σ [$F(1, 38) = 1.304$, $p=0.261$] in nonword targets. These findings were converging with the analysis in deciles, apart from the absence of significant effects in word targets.

5.2.3 Discussion

The experiment replicated the response congruency effect that was reported in Experiment 2. This experiment aimed to replicate the null effect reported in the literature (Norris & Kinoshita, 2008; Perea et al., 1998, 2010) by replacing the nonword targets used in Experiment 2 with items that increased the difficulty of the lexical decision task. The effect size in this experiment was still large at around $f=0.5$, but it was considerably smaller than in Experiment 2 where it was approximately $f=1.0$. This suggested that the manipulation was effective as measured in a reduced effect size and a smaller congruency effect in absolute numbers. But the manipulation was not sufficiently strong to reduce the effect to a null. This finding indicated that a stronger manipulation might explain the empirical difference.

Table 5.2: Results of the Ex-Gaussian fitting using QMPE as a function of target lexicality and congruency of Experiment 3.

	Word			Nonword		
Parameter	Incongruent	Congruent	*Effect*	Incongruent	Congruent	*Effect*
μ	442	431	*11*	482	460	*22*
σ	43	46	*-3*	40	46	*-6*
τ	72	66	*6*	77	91	*-14*

Looking at the nonword targets, there was a response congruency effect in the repeated measures analysis. On closer examination, using the analysis by deciles this effect was prevalent in the faster responses. In slower responses, the congruency effect turned into a negative effect; that is incongruent trials were responded to faster than congruent trials. This was further supported by the Ex-Gaussian analysis, showing a significant effect on the exponential part of the RT distribution. Remarkably, there was still a strong positive congruency effect in the central tendency of the distribution. Replacing the nonword targets reduced the effect size and resulted in some trials showing a negative effect. Thus, the null effect reported in the literature could be due to a mixture of positive and negative effects. This would suggest that the distinctive difference between the experiments is the task difficulty and it implied that increasing the difficulty further could eliminate the congruency effect in mean RT.

The SCM (Davis, 2010) as well as the leaky accumulator model (Usher & McClelland, 2001) suggest that the effect is decreasing with slower reaction times. The results of this experiment showed that the response congruency effect was reduced in the mean RT with more difficult nonword targets. This contrasts with reports of an increasing semantic effect in τ by Balota et al. (2008), whereby the exponential part of the distribution was larger in the congruent condition and the priming effect increased with the later deciles. Also, the finding contrasts with Bodner and Masson (1997) who suggested that the prime would increase its effect on more effortful processing. If the task difficulty of the explicit task determines whether a response congruency effect is present, this would support predictions by the SCM and the leaky accumulator model. But it is important to note that neither of these models predicts that the effect turns into a negative effect. The following experiment tested whether an increased task difficulty is the underlying reason for the presence or absence of a congruency effect, but it also tested whether a negative effect can be replicated.

5.3 Experiment 4

This experiment was designed to examine whether the prime congruency effect is affected by manipulating the word targets.

Table 5.3: Mean reaction times in ms and error ratio in percent as a function of lexicality and congruency of Experiment 4.

	Lexicality			
	Word		Nonword	
Congruency	RT	Error	RT	Error
Incongruent	541	3.98	550	1.29
Congruent	529	3.15	538	1.59
Effect	12	0.83	12	-0.30

5.3.1 Methods

Participants. Thirty-five participants from the same population as in Experiment 1 took part in the study.

Stimuli & Design. The same nonwords as in Experiment 2 were used as primes and targets, the word primes were also kept constant. The word targets were exchanged for words of a lower frequency reflecting a medium range (mean=41.19/million, min=0/million, max=432.91/million; Baayen et al., 1995; N-Watch results). These words also scored medium on a typicality measure ($OT3_{mean}$=0.20, $OT3_{max}$=1.31, $OT3_{min}$=-1.55; see Chapter 2.3), e.g., *scoop*, *venue*. Again, all stimuli were formed of five letters. Primes and targets were paired so that there were no letters shared between prime and target (see Appendix B for a list of all stimuli). Two versions of the experiment were formed for counterbalancing purposes.

Procedure. The procedure was the same as in Experiment 2.

5.3.2 Results

One participant showed extremely slow RTs with a mean of 1366 ms and was excluded from further analysis. As in previous experiments, participants and items showing an error rate greater than 25% were dropped from the analysis. This affected no participants, but one item (*bison*). Half of the participants were assigned to list 1 and half to list 2. Outliers were removed by 3 SDs for each participant using the correct responses. This affected 1.77% of the remaining data. The results are presented in Table 5.3.

Repeated measures analysis of variance

An ANOVA was performed using the participant-wise *z*-scores. The list factor entered all analyses as a between participants factor and congruency as a repeated within participants factor. Lexicality was a repeated factor within participants in F_1 and a between items factor in F_2. There was a tendency of a main effect of lexical state in *z*-scores [$F_1(1, 32) = 2.901$, $p=0.098$; $F_2(1, 195) = 3.730$, $p=0.055$] indicating that word stimuli were responded to slightly faster than the nonword stimuli.

In word targets there was a main effect of congruency on *z*-scores [$F_1(1, 32) = 21.733$, $p<0.001$, $f=0.82$; $F_2(1, 97) = 20.264$, $p<0.001$, $f=0.46$] showing that congruent word trials received faster responses than incongruent word trials. Similarly, there was a main effect in nonword targets [$F_1(1, 32) = 21.282$, $p<0.001$, $f=0.81$; $F_2(1, 98) = 12.861$, $p=0.001$, $f=0.36$]. This indicated that congruent nonword trials were responded faster than incongruent trials.

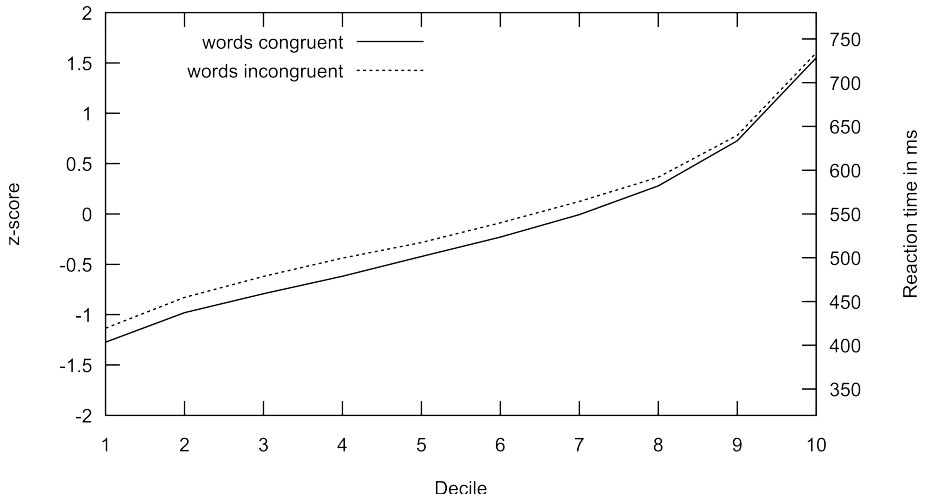

Figure 5.2: Plot of reaction times in word targets of Experiment 4 by deciles. The scale of reaction times is approximation to reflect the actual corresponding z-score and reaction time.

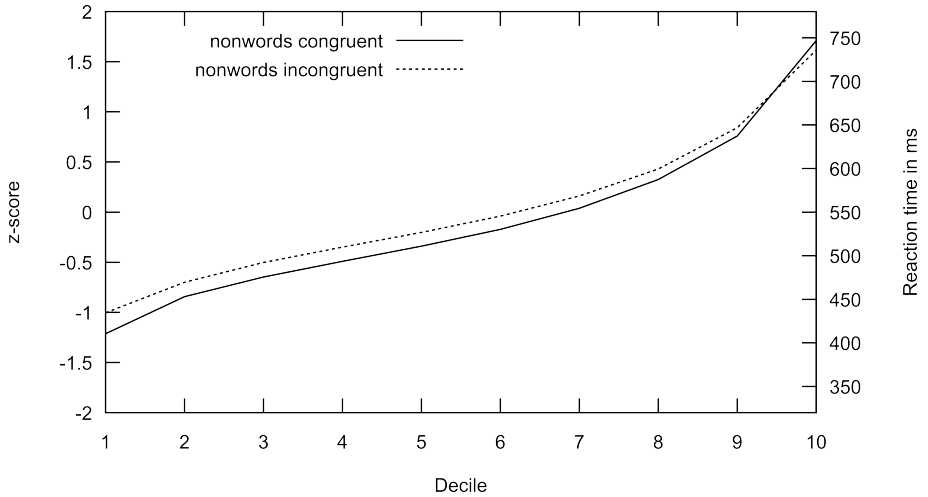

Figure 5.3: Plot of reaction times in nonword targets of Experiment 4 by deciles. The scale of reaction times is approximation to reflect the actual corresponding z-score and reaction time.

There was no main effect of congruency on the error ratio in words [$F_1(1, 32) = 1.876$, $p=0.180$; $F_2(1, 97) = 1.505$, $p=0.223$] or nonwords [$F_1(1, 32) = 0.378$, $p=0.543$; $F_2(1, 98) = 0.462$, $p=0.480$].

Analysis in deciles

The reaction times were binned in ten vincentiles and the results submitted to an analysis using lme4 in R (see Appendix E for a detailed list of results). With regards to word targets, the results showed that the congruency effect was significant in all deciles apart from the slowest two. This was supported by the visual inspection of the RT distribution in Figure 5.2, where the advantage of congruent over incongruent word trials declined towards the slower responses. In nonword targets the congruency effect

was significant in all deciles. But it is important to note that the sign of the *t*-value in the last decile is negative and the lines in Figure 5.3 cross. That means in fast responses the congruent trials showed an advantage over incongruent trials, but in very slow reaction times the response congruency effect turned into a negative effect. A similar pattern was observed in Experiment 3.

Ex-Gaussian analysis

The fitting of the Ex-Gaussian parameters was performed according to the method outlined in Experiment 1 and results are shown in Table 5.4. There was a significant main effect of congruency in word targets on μ [$F(1, 32) = 20.534$, $p<0.001$, $f=0.80$] and on τ [$F(1, 32) = 6.145$, $p=0.019$, $f=0.44$]. There was no effect in σ [$F(1, 32) = 2.297$, $p=0.139$]. The effect in μ reflects an advantage of congruent over incongruent trials in the central tendency of the RT distribution. In contrast, the effect in τ reflects a greater number of very slow responses in congruent trials compared to incongruent trials. The analysis in nonword targets showed a significant effect of congruency on μ [$F(1, 32) = 10.938$, $p=0.002$, $f=0.59$] indicating faster responses in congruent trials. There was no effect on σ [$F(1, 32) = 0.046$, $p=0.832$] or τ [$F(1, 32) = 1.293$, $p=0.264$]. Even though there was a negative congruency effect in the last decile the effect in τ was not significant. The difference in word and nonword targets with respect to τ reflects a smaller difference in very slow responses between congruent and incongruent trials in nonword targets despite the absence of a negative effect in words.

5.3.3 Discussion

In this experiment a response congruency effect was found in both word and nonword targets. The word targets were exchanged for words with a lower frequency

Table 5.4: Results of the Ex-Gaussian fitting using QMPE as a function of target lexicality and congruency of Experiment 4.

	Word			Nonword		
Parameter	Incongruent	Congruent	*Effect*	Incongruent	Congruent	*Effect*
μ	467	441	*26*	473	454	*19*
σ	44	36	*8*	37	38	*-1*
τ	74	91	*-17*	78	85	*-7*

and a lower score in orthographic typicality compared to Experiment 2. The manipulation resulted in a reduced effect size compared to Experiment 2, which in total was similar to Experiment 3. This indicated that the manipulation was effective, but replacing either the nonword targets or the word targets was not a sufficiently strong manipulation to eliminate the congruency effect. Though, the reduced effect sizes suggested that combining harder nonword and word targets may result in a null effect and provide an explanation for the discrepancy with the literature (Norris & Kinoshita, 2008; Perea et al., 1998, 2010).

Interestingly, in Experiment 3 as well as in Experiment 4 the congruency effect diminished or turned into a negative effect in slow responses. Furthermore, there was a significant effect in the exponential component of word targets in Experiment 4. This suggested that such a negative effect is more likely the slower the responses in an experiment are. The null effects reported in the literature (Norris & Kinoshita, 2008; Perea et al., 1998, 2010) could be due to a mixture of positive and negative effects or a very short-lived effect. Both could cover a response congruency effect in mean RT, but in both cases the congruency effect was not absent, rather it was not detectable by analysing mean RTs. Thus, the null effects reported in the literature could be cases where an effect remained undiscovered.

In lexical decision tasks word targets typically receive faster responses than the nonword targets, but in this experiment the lexicality effect did not reach significance. The reduced effect of lexicality substantiated that the specific selection of word targets showed the expected slow down in word responses. This did not affect the response speed to nonword targets as strongly and hence the lexicality effect was attenuated.

The next experiment tested whether the effect diminishes if hard word and nonword targets are used and also whether there is an increase in the negative priming effect.

5.4 Experiment 5

In this experiment the two factors that increased the difficulty of the lexical decision task in the previous two experiments were combined. By combining the more difficult word targets of Experiment 4 and the more difficult nonword targets of Experiment 3, the degree of difficulty of the task was increased above the levels of the previous experiments. The previous experiments indicated that increasing the difficulty of the task results in attenuated congruency effects. Thus, the current experiment could show a null effect of response congruency. By showing that the response congruency effect diminished in this set of targets the presence or absence of the congruency effect could be attributed to the task difficulty.

5.4.1 Methods

Participants. Twenty-four participants from the same population as in Experiment 1 took part in the experiment.

Stimuli & Design. All primes in this experiment were the same as in Experiment 2. The nonword targets were taken from Experiment 3 and the word targets from Experiment 4. Primes and targets were paired so that there were no letters shared between prime and target (see Appendix B for a list of all stimuli). Two versions of the experiment were formed for counterbalancing purposes.

Procedure. The procedure was the same as in Experiment 2.

5.4.2 Results

As in previous experiments, participants and items showing an error rate greater than 25% were dropped from the analysis. This affected no participants and two items (*bison*; *avail*). Every list was assigned half of the participants, resulting in 12 per list. Outliers were removed by 3 SDs for each participant using the correct responses. This affected 1.39% of the remaining data. The results are presented in Table 5.5.

Table 5.5: Mean reaction times in ms and error ratio in percent as a function of lexicality and congruency of Experiment 5.

	Lexicality			
	Word		Nonword	
Congruency	RT	Error	RT	Error
Incongruent	539	3.57	586	2.75
Congruent	531	4.08	576	2.83
Effect	8	-0.51	10	-0.08

Repeated measures analysis of variance

ANOVAs were performed using the participant-wise *z*-scores of correct responses. The list factor was always added as between participants factor and congruency entered the analysis as repeated within participants factor. Lexicality was a repeated within factor in F_1 and a between items factor in F_2. There was a main effect of lexicality in *z*-scores [$F_1(1, 22) = 22.083$, $p<0.001$, $f=1.00$; $F_2(1, 194) = 76.001$, $p<0.001$, $f=0.63$] reflecting that word stimuli received faster responses than nonwords in this experiment. All further analyses were performed individually for word and nonword targets.

In word targets, the results showed a significant effect of congruency in *z*-scores [$F_1(1, 22) = 4.794$, $p=0.039$, $f=0.47$; $F_2(1, 96) = 5.893$, $p=0.017$, $f=0.25$] indicating that congruent word trials received faster responses than incongruent word trials. In

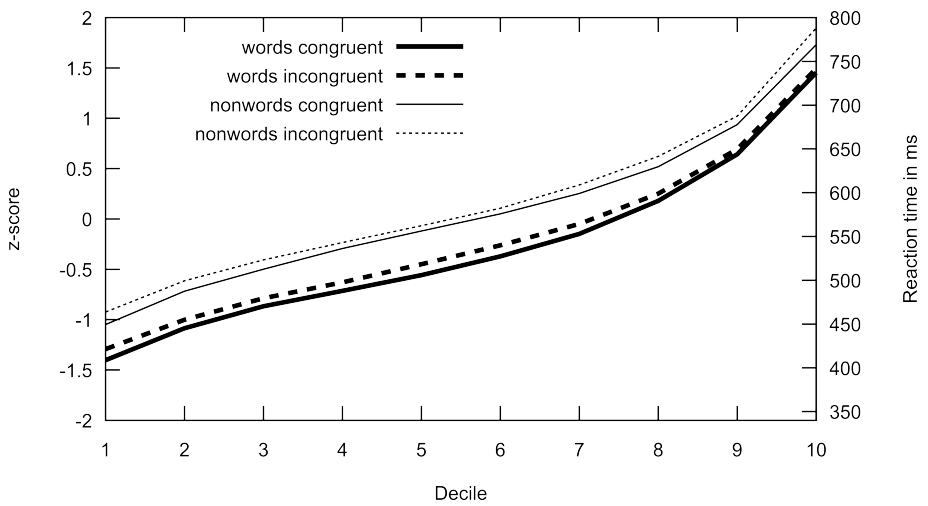

Figure 5.4: Plot of the results of Experiment 5 by decile and condition. The scale of reaction times is an approximation to reflect the actual corresponding value of z-score and reaction times.

nonword targets, there was a significant effect of congruency in *z*-scores [$F_1(1, 22) = 5.863$, $p=0.024$, $f=0.52$; $F_2(1, 98) = 5.525$, $p=0.021$, $f=0.24$] showing that congruent nonword trials were responded faster than incongruent nonword trials.

The analysis of error ratio in word targets did not show a significant effect of congruency [$F_1(1, 22) = 0.354$, $p=0.588$; $F_2(1, 96) = 0.366$, $p=0.546$] and there was no effect in nonword targets either [$F_1(1, 22) = 0.031$, $p=0.863$; $F_2(1, 98) = 0.016$, $p=0.901$].

Analysis in deciles

The *z*-score transformed reaction times were binned in ten vincentiles and submitted to an analysis using lme4 in R (see Appendix E for a detailed list of results). In word targets the congruency effect was significant in the first eight deciles indicating faster responses to congruent word trials than to incongruent trials. But the effect was not significant in the two slowest bins, which was compatible with the visual inspection of Figure 5.4. In nonword targets, there was a significant positive effect of congruency in fast as well as in slow reaction times. However, the effect was not significant in the centre of the distribution in deciles five and six.

Ex-Gaussian analysis

The RT data were fitted an Ex-Gaussian distribution as outlined in Experiment 1 and the results are shown in Table 5.6. There was a significant congruency effect in word targets in μ [$F(1, 22) = 6.932$, $p=0.015$, $f=0.56$] indicating faster responses to congruent trials than to incongruent trials. There was no effect in σ [$F(1, 22) = 0.036$, $p=0.852$] or τ [$F(1, 22) = 0.058$, $p=0.812$]. The analysis in nonword targets did not reveal significant

Table 5.6: Results of the Ex-Gaussian fitting using QMPE as a function of target lexicality and congruency of Experiment 5.

	Word			Nonword		
Parameter	Incongruent	Congruent	*Effect*	Incongruent	Congruent	*Effect*
μ	455	445	*10*	510	502	*8*
σ	36	35	*1*	46	51	*-5*
τ	86	87	*-1*	73	71	*2*

effects of congruency in μ [$F(1, 22) = 0.952$, $p=0.340$], σ [$F(1, 22) = 0.364$, $p=0.552$] or τ [$F(1, 22) = 0.028$, $p=0.868$].

5.4.3 Discussion

The results of this experiment showed a prime congruency effect. In word targets, the central tendency was shifted showing an advantage of congruent trials. This was supported in all analyses. In nonword targets the analysis in repeated measures as well as by deciles showed an effect, whereas the parameters derived from an Ex-Gaussian distribution did not. The latter could be due to the smaller effect in the central deciles of the RT distribution. Comparing the numerical effect in this experiment of about 8 ms in words and 10 ms in nonwords and the effect found in Experiment 2 of about 23 ms, the nominal effect size decreased clearly. A similar pattern was found in the effect size which was smaller than in the previous experiments. This demonstrated that with increasing task difficulty the congruency effect decreased. However, it also showed that the congruency effect is robust even in difficult tasks. That means the experiment was not yet compatible with the null effect reported in the literature.

A negative response congruency effect did not occur in nonword targets, which is in contrast to Experiment 3 and Experiment 4. Also, there was no significant effect of congruency in the exponential part of the RT distribution. In fact, the estimates of τ were very similar across conditions. This suggested that a negative congruency effect is not dependent on the task difficulty per se, but an asymmetry in the difficulty of the word and nonword targets as in Experiment 3 and 3 could contribute to their occurrence. The experiment showed that the task difficulty contributes to the presence and absence of a congruency effect, but it does not determine it.

5.5 General discussion

The three experiments in this chapter tested the robustness of the congruency effect established in Experiment 2 with regards to task difficulty. The predictions derived from computational models such as the Spatial Coding Model (Davis, 2010) and the leaky accumulator model (Usher & McClelland, 2001) suggested that priming effects could be reduced in magnitude in slower responses. This is due to averaging

evidence across cycles in the SCM's decision component and trickling out of the priming effect respectively. The mean reaction times of all experiments supported this. Comparing experiments Experiment 3 and 4 to Experiment 2, the effect size and the numerical size of the congruency effect decreased with increased task difficulty. Furthermore, comparing Experiment 5 to the three experiments the effect decreased even further. These results challenge the prediction by Bodner and Masson (1997). Apparently, the response congruency effect did not increase with an increase of opportunity for the prime to affect the processing.

5.5.1 Distributional analysis

A more fine grained analysis of the reaction times revealed that the response congruency effect was not only reduced, but it turned into a negative effect. A mixture of positive and negative effects could result in the null effects that were previously reported (Norris & Kinoshita, 2008; Perea et al., 1998, 2010). This would suggest that a priming effect was not absent, but rather that it was not detected.

In Experiment 3, the nonword targets were harder to reject than in Experiment 2 and this resulted in a significant effect on the exponential part of the RT distribution reflected in τ in the Ex-Gaussian fit. In Experiment 4, the word targets were replaced and a similar effect was found in τ. In Experiment 3 and 4 the central tendency μ in the category with an increased difficulty showed a numerical difference of about 20 ms between congruent and incongruent trials. In these items τ showed an effect in opposite direction indicating that the number of very slow responses was greater in the congruent trials. With respect to mean reaction times, these two parameters could cancel each other out and a null effect would be reported. Remarkably, the negative congruency effect did not occur in Experiment 5 where both word and nonword targets were of a similar level of difficulty. In summary, a negative congruency effect and the corresponding effect in τ occurred in the two experiments (3 and 4) which used an asymmetric level of difficulty in the response categories. In contrast, no such an effect occurred in both experiments with a balanced difficulty of the task (2 and 5).

5.5.2 Negative congruency effects

The presence of positive and negative effects is not equivalent to the absence of an effect, even though the mean RT could suggest this. In a number of experiments where participants had to indicate the orientation of a horizontal line, Boy and Sumner (2010) showed that negative priming effects only occur when positive priming effects could occur. In their experiments a short SOA (40 ms) triggered positive effects whereas longer SOAs (150 ms) triggered negative priming effects. This pattern was attributed to self-inhibition. In this model (Boy & Sumner, 2010) each response option is subject to a response threshold and an inhibition threshold. If the response threshold is hit, an explicit motor response is triggered. This usually occurs due to the target stimulus. If the inhibition threshold has not been hit, the system tries to deal with the activation as noise. This is the source of positive priming effects. Once the activation hits the inhibition threshold, it is recognised and is 'eligible' for inhibition (see Colombo, 1986 for a similar suggestion). This mechanism predicts an advantage in incongruent cases, where the activation in the prime response hit the inhibition threshold but not the response threshold. In these cases, active suppression of the incongruent response provides an advantage for the correct target response. In congruent trials this suppression results in a disadvantage, because the actual target response is suppressed. In Boy and Sumner's (2010) experiments, the long SOA allowed the primes to hit the inhibition threshold and this produced negative effects, whereas shorter SOAs did not allow the prime to hit the threshold (see Eimer & Schlaghecken, 2003 for similar suggestions and review; see Kiesel, Berner, & Kunde, 2008 for mask induced account). In this account, negative priming effects are essentially due to the same mechanism as positive effects and an additional self-inhibition process. This means that negative effects can only occur if there is at least a potential for positive effects.

An inhibitory effect in the decision channels could also result from other sources. The activation of a prime could have been passed on to a semantic level, where an image or sensation could be in mismatch with the stimulus. The decision channels appear to be open to influence from a semantic level in general. Examples include valence (Kousta et al., 2009), imageability (Balota, Cortese, Sergent-Marshall, Spieler, &

Yap, 2004) and concreteness (Kroll & Merves, 1986; Schwanenflugel, Kipp Harnishfeger, & Stowe, 1988). Nevertheless, empirical data from semantic priming effects suggests that there is (possibly) a negative effect in μ and a positive in τ (Balota et al., 2008). This means that the central tendency of the distribution indicates that congruent cases receive slightly slower responses, whereas the number of very slow responses is greater in incongruent compared to congruent cases. Although, there were significant effects in τ, in all four experiments presented so far there was no sign of the congruency effect becoming larger in slow reaction times. In contrast, there were positive effects in μ and negative effects in τ. These differences suggest that the congruency effect is not a semantic effect.

A verification mechanism is also capable of explaining the data pattern reported in the present experiments. In contrast to self-inhibition which operates on response level, a verification mechanism operates on word nodes. Grainger and Jacobs (1999) argued in favour of a reset in word nodes that was similar to the adaptive resonance theory (Carpenter & Grossberg, 1987) and the activation-verification model (Paap, Newsome, McDonald, & Schvaneveldt, 1982). This mechanism resets a word node if it mismatches the stimulus. A word prime triggers activation in the lexicon and contributes to summed lexical activity. A nonword target could activate some word nodes that are similar to the nonword and contribute to summed lexical activity as well. Once the lexical activity has hit a threshold the verification mechanism is initiated. In a nonword trial the most active word node produces a mismatch and undergoes a reset. As a result the correct *no*-response is triggered faster than in case of a nonword prime that did not contribute to the summed lexical activity. That means the nonword response was triggered faster in the incongruent trial than in the congruent trial. A verification process could also result in negative congruency effects in word trials. If the word target is similar to another word which is higher in frequency, this similar node could be strongly activated. A word prime and the competitor word node would contribute to summed lexical activity and trigger the verification. In this case a reset would slow a correct *yes*-response on the basis of summed lexical activity and force the model to wait for the identification of the word target. In an incongruent word trial,

the nonword prime would contribute only little to summed lexical activity and a verification is less likely to result in a mismatch.

Positive congruency effects were observed in most trials and thus, the threshold for triggering a verification has to be set sufficiently high. The current experiments suggested that this threshold should be sensible to the respective task, because these effects were most prevalent in slow responses in experiments when one of the categories was more difficult than the other. Grainger and Jacobs (1996) suggested a variable deadline model that provided a mechanism for establishing at what value of summed lexical activation (Σ-criterion) a response should be triggered. In contrast, the verification threshold could be fixed. This implies that in clear cases, a response is triggered immediately, but in more difficult trials a verification is initiated. In particular, this would expose the category that was manipulated in Experiment 3 (nonwords) and 4 (words) to be subject to a verification. Thus, the congruency effect and the anti-congruency effect would depend on summed lexical activity. This is compatible with the empirical data. As shown in Chapter 4.2, the SCM (Davis, 2010) captured the response congruency effect in Experiment 2 through an increased level of summed lexical activity in the early target related cycles in congruent compared to incongruent word trials and vice versa in nonword trials. That means the congruency effect was essentially generated by a greater activity in the *yes*-channel in congruent compared to incongruent word trials and vice versa in nonwords. Thus, a reset in the word node that mainly contributes to summed lexical activity appears to be a plausible option for producing negative congruency effects. I return to this point in Chapter 8 when presenting the respective simulations.

Bodner and Masson's (1997) hypothesis that the prime increases its effect with more effortful processing was not incompatible with the data. However, the mixture of positive and negative effects could prevent the detection of an increasing effect, i.e. the effect size could be continuously growing but the onset of negative effects covers this. The presence of negative effects suggested that the effect of the prime is different in fast and slow responses. Once the prime was processed up to certain level its effect is inhibitory. If the pattern of negative congruency effects is replicable and a larger

proportion of the distribution shows a negative effect, computational models will be challenged.

5.5.3 Theoretical accounts of response congruency

The results of Experiment 2 to 4 provided evidence in favour of a model that can explain response congruency effects independent of the difficulty of the explicit task. The deep processing account (Dehaene et al., 1998) predicts that response congruency effects are dependent on the prime and its processing, but not on the target or the difficulty of the explicit task (see 3.4.2). The current results provided support for this assumption. The deep processing account also predicts that response congruency effects are due to interference at a motor level, more precisely positive effects were attributed to suppressing an erroneous readiness potential. If the negative effects that were observed in Experiment 3 and 3 can be explained by the self-inhibition account, there is additional evidence that these effects were the result of interference at response readiness level.

The semantic overlap model (Quinn & Kinoshita, 2008) predicts response congruency effects independently of the difficulty of the explicit task as well. This model attributed the effect to features that were activated and attributed the benefit in responses to an overlap in congruent cases. Other empirical data (Balota et al., 2010, 2008) suggested that semantic effects are accompanied with increasing effect sizes towards the slow tail of the RT distribution. On the other hand, a similar mechanism to verification could be operational in semantic features. If a feature that is part of the overlap between prime and target was reset, a disadvantage in the congruent condition was predicted. In the incongruent condition a reset of a feature would not cause a slow down in the response.

Both, the deep processing account and the semantic feature overlap hypothesis predict that response congruency crucially depends on the properties of the prime. If the prime was hard to process the readiness potential could be smaller and the deep processing account would predict the absence of response congruency effects. In the feature overlap model a response congruency effect is not predicted if the primes

would not show overlapping features with the targets of either category or if the primes of both categories would share features with both target categories. The experiments in the next chapter tested whether the response congruency effect is dependent on the primes.

5.5.4 Empirical discrepancy

The experiments in this chapter clearly demonstrated response congruency effects in lexical decision. This is compatible with one of the experiments in the literature (Grainger & Jacobs, 1996), but it is in contrast to other studies (Norris & Kinoshita, 2008; Perea et al., 1998, 2010). First, all experiments presented so far showed that congruency effects can occur even if there is no form relation or shared semantics between prime and target. Secondly, the manipulation of the task difficulty suggested that in more difficult tasks with slower responses, the effect size decreases. One option is to attribute the absence of congruency effects to the emergence of negative effects in slow responses, but all experiments showed a significant positive congruency effect in mean RT despite the negative effects in slow responses. In all experiments the same set of very informative primes was used. If the prime was less informative with regards to the task, the congruency effect could diminish. This would provide information about the sort of information that was extracted from the prime. In the following chapter I introduce experiments that tested this prediction, before continuing with simulating the current findings.

5.6 Conclusion

A response congruency effect was found in lexical decision experiments with an increased task difficulty. Though the effect was significant in both word and nonword targets in four experiments, the effect size decreased with an increased task difficulty. This suggested that the targets contribute to the effect, but the difficulty of the explicit task does not determine the presence or absence of congruency effects. The effect was not short-lived and was still present in slower reaction times. Although the congruency effect can turn into a negative effect in nonword targets as a more fine grained analysis of the RT distribution revealed. Computational models, such as the Spatial Coding Model (Davis, 2010) predict decreasing effect sizes with slower responses and are

compatible with the empirical data. The next chapter investigates the influence of the primes.

6. Effects of prime difficulty

The experiments in the previous chapters tested whether the response congruency effect is affected by manipulations of the explicit task. All experiments showed the presence of a response congruency effect, even when the difficulty of the task was increased and slower responses were observed. However, one important factor was held constant in these experiments was the informativeness of the primes. This chapter introduces experiments that manipulated this factor. I begin by introducing the concept of prime informativeness.

6.1 Prime informativeness

All primes used in Experiment 2 through 5 had the property of being either extremely wordlike or extremely unwordlike. The word primes were high in frequency and resembled typical letter combinations (e.g., *night*, *youth*, *crown*; more details on the exact computation of typicality are provided in Chapter 2.3). Thus, these words trigger a comparably high level of lexical activation. By contrast, the nonwords were very unlike English words: they had no orthographic neighbours and were formed of atypical letter strings (e.g., *cxnio*, *dvnel*, *sujcw*). These nonwords trigger very little activation in the lexicon. Mapping them on a scale of how likely each item is to be a word, the words and nonwords would be very distant. In this respect, the primes can be described as very informative with regards to a lexical decision task. These informative primes could have a greater impact on the final decision than primes that are closer and more similar on a wordlikeness scale. Previous experiments (Norris & Kinoshita, 2008; Perea et al., 1998, 2010) that did not observe response congruency effects used primes that were similar to the targets with respect to wordlikeness. Thus, the presence of congruency response effects in Experiment 2 through 4 could have been critically dependent on the nature of the primes. This possibility was tested in the experiments reported in this chapter.

6.2 Experiment 6

This experiment combined the relatively difficult targets from Experiment 5 with primes that were similarly difficult, i.e., relatively uninformative. Hence, if prime

informativeness is a critical factor the congruency effect (which was relatively small in Experiment 5) could be reduced to a null effect. This would be compatible to the literature and shed light on the crucial difference between this and the previous experiments.

6.2.1 Methods

Participants. Twenty-five participants from the same population as in Experiment 1 took part in the experiment. All were native speakers of English.

Stimuli & Design. 100 word targets were taken from Experiment 5 and were the same as in Experiment 4. The 100 nonword foils were also mostly taken from Experiment 5 and therefore virtually the same as in Experiment 3. Five of the nonwords had to be replaced from the set used in Experiment 3 and 5 in order to fulfil restrictions in forming the pairs of nonwords in the congruent condition ($OT3_{mean}$=-0.52, $OT3_{max}$=0.51, $OT3_{min}$=-3.41; for a full list see Appendix B). The critical difference from the previous experiments was the use of primes that were relatively less informative. The word and nonword primes were selected by sampling without replacement from the set of targets, re-pairing primes with targets in such a way that primes and targets shared no letters. For counterbalancing purposes, two versions of the experiment were derived.

Procedure. The procedure was the same as in Experiment 2.

6.2.2 Results

One participant showed exceptionally slow RTs (mean of 1109 ms) and was therefore excluded from further analysis. As in previous experiments, participants and items showing an error rate greater than 25% were dropped from the analysis. This criterion did not affect participants, but three word targets (*motto*, *bison*, *avail*) were excluded. Half of the participants was assigned to each list. Outliers were removed by 3 SDs for each participant using the correct responses. This affected 1.80% of the remaining data. The mean reaction times and the error ratios are presented in Table 6.1.

Table 6.1: Mean reaction times in ms and error ratio in percent as a function of lexicality and congruency of Experiment 6.

	Lexicality			
	Word		Nonword	
Congruency	RT	Error	RT	Error
Incongruent	556	4.30	593	2.75
Congruent	542	4.04	596	2.83
Effect	14	0.26	-3	-0.08

Repeated measures analysis of variance

ANOVAs were performed using the participant-wise *z*-scores. In all analyses the list factor entered as a between participants factor and congruency was added as repeated within participants factor. Lexicality was a repeated within participants factor in F_1 and a between items factor in F_2. There was a main effect of lexical status in *z*-scores [$F_1(1, 22) = 34.408$, $p<0.001$, $f=1.25$; $F_2(1, 193) = 66.539$, $p<0.001$, $f=0.59$] reflecting that word stimuli received faster responses than nonwords in this experiment. All further analyses treated word and nonword targets individually.

In word targets the congruency effect was significant in *z*-scores [$F_1(1, 22) = 7.731$, $p=0.011$, $f=0.59$; $F_2(1, 95) = 11.572$, $p=0.001$, $f=0.35$] indicating faster responses to congruent than to incongruent trials. In contrast, there was no such effect in nonword targets [$F_1(1, 22) = 0.020$, $p=0.890$; $F_2(1, 98) = 0.095$, $p=0.759$].

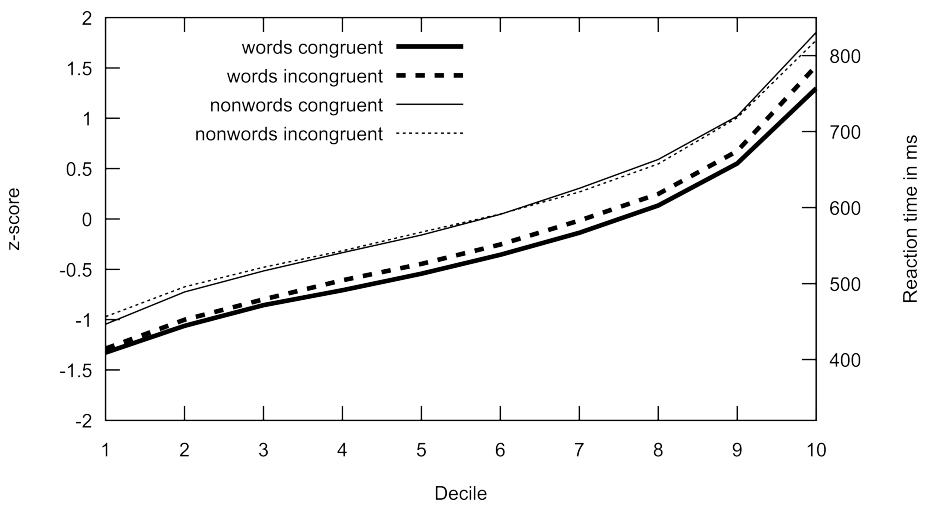

Figure 6.1: Plot of the results of Experiment 6 by decile and condition. Note that the scale in ms in an approximation and data was plotted by z-scores.

The analysis of error ratio in word targets did not reveal a main effect of congruency [$F_1(1, 22) = 0.066$, $p=0.800$; $F_2(1, 95) = 0.087$, $p=0.769$]. This was also the case in nonword targets [$F_1(1, 22) = 0.203$, $p=0.656$; $F_2(1, 98) = 0.181$, $p=0.671$].

Analysis in deciles

The reaction times in *z*-scores were binned in ten vincentiles before being submitted to an analysis using lme4 in R (see Appendix E for a detailed list of results). In word targets the congruency effect was significant in all but the first three deciles. This was compatible with the visual inspection of Figure 6.1, where the effect was increasing with increasing RT. Also, the *t*-values were increasing in greater deciles indicating a greater effect size. In nonword targets there was a significant positive congruency effect in the first decile and a significant negative effect in the last decile. The *t*-values indicated that the tendency for a positive effect became smaller and the *t*-value turned negative in decile seven, indicating a negative congruency effect in slower responses. Figure 6.1 illustrates that the RT curves of congruent and incongruent nonword targets were crossing. The effect was small in all deciles.

Ex-Gaussian analysis

There were no significant effects in the parameters of the Ex-Gaussian analysis. There was no main effect of congruency in word targets on μ [$F(1, 22) = 1.893$, $p=0.183$], on σ [$F(1, 22) = 0.992$, $p=0.330$] or τ [$F(1, 22) = 0.546$, $p=0.468$]. Similar results were found in nonword targets on μ [$F(1, 22) = 0.272$, $p=0.607$], σ [$F(1, 22) = 0.660$, $p=0.425$] or τ [$F(1, 22) = 0.799$, $p=0.381$].

Table 6.2: Results of the Ex-Gaussian fitting of Experiment 6 using QMPE as a function of target lexicality and congruency.

	Word			Nonword		
Parameter	Incongruent	Congruent	*Effect*	Incongruent	Congruent	*Effect*
μ	460	451	*9*	500	496	*4*
σ	47	41	*6*	44	49	*-5*
τ	97	92	*5*	97	103	*-6*

6.2.3 Discussion

The results showed that there was a prime congruency effect in word targets, but not in nonword targets. These results provided some support for the role of prime informativeness. The response congruency effect that was observed in Experiment 5 disappeared when the same nonword targets were paired with less informative primes in the present experiment. This manipulation did not the reduce the effect size in word targets. But the congruency effect was detected in mean RT only, whereas the central tendency, standard deviation and exponential part of the RT distribution did not indicate a significant effect. Thus, manipulating the prime informativeness affected the word responses in this experiment. In summary, the response congruency effect was decreased by reducing the prime informativeness compared to Experiment 5.

Five separate experiments have provided evidence for a significant response congruency effect in word targets, and four out of these have shown evidence for a significant response congruency effect in nonword targets. These effects appear to be very robust and are yet at odds with prior findings in the literature (Norris & Kinoshita, 2008; Perea et al., 1998, 2010). However, the experiments reported so far have shown that the size of these effects is influenced by the difficulty of the word-nonword discrimination task. The experiments in the last chapter showed that congruency effects can be reduced by making the target stimulus discrimination more difficult, and the present experiment showed that combining relatively difficult targets with relatively difficult (uninformative) primes can eliminate the congruency effect in nonword targets. This suggested that making the word-nonword discrimination even more difficult in primes and targets could completely eliminate the congruency effect. This possibility was tested in Experiment 7.

6.3 Experiment 7

The word and nonword targets in this experiment were selected so as to maximise the difficulty level of the explicit task. The same stimuli were also employed as primes, so that the prime informativeness was also reduced relative to the previous experiment. Since word targets were already selected in Experiment 6 to be relatively

hard to categorise, only nonwords could be manipulated further to achieve an increased task difficulty, without losing considerable amounts of data due to erroneous responses. Hence, the nonword primes and targets were replaced by more wordlike stimuli.

6.3.1 Methods

Participants. Thirty-six participants from the same population as in Experiment 1 took part in the experiment.

Stimuli & Design. The stimuli in this experiment were selected to maximise the difficulty of the task. The 100 low frequency words used in Experiments 3, 4 and 5 were used as primes and targets. The words in this experiment had 1.5 neighbours on average according to N-watch (Davis, 2005). The nonwords were replaced by 100 items that had 3.7 neighbours on average according to N-Watch and were higher in orthographic typicality ($OT3_{mean}$=1.48, $OT3_{max}$=1.77, $OT3_{min}$=1.33; see Chapter 2.3) than the words ($OT3_{mean}$=0.20, $OT3_{max}$=1.31, $OT3_{min}$=-1.55; see Chapter 2.3). These nonwords were used as primes and targets (see Appendix B). As in the previous experiments, the primes and targets were paired so that there were no letters shared between prime and target. There were two versions of the experiment for counterbalancing purposes.

Procedure. The procedure was the same as in Experiment 2.

6.3.2 Results

As in previous experiments, participants and items showing an error rate greater than 25% were dropped from the analysis. This did not affect participants, but two words (*inert*, *bison*) and eight nonwords (*wence*, *goven*, *becon*, *traid*, *quess*, *drome*, *sping*, *hince*) were excluded. The larger number of exclusions indicated that the difficulty of the explicit lexical decision task was increased, as expected. Outliers were removed by 3 SDs for each participant using the correct responses. This affected 1.67% of the remaining data. The results are presented in Table 6.3.

Table 6.3: Mean reaction times in ms and error ratio in percent as a function of lexicality and congruency of Experiment 7.

	Lexicality			
	Word		Nonword	
Congruency	RT	Error	RT	Error
Incongruent	638	3.68	794	6.75
Congruent	636	3.57	793	9.98
Effect	2	0.11	1	-3.23

Repeated measures analysis of variance

ANOVAs were performed using the participant-wise *z*-scores. In all analyses the list factor entered as a between participants factor and congruency was added as repeated within participants factor. Lexicality was a repeated within participants factor in F_1 and a between items factor in F_2. There was a main effect of lexicality in *z*-scores [$F_1(1, 34) = 263.507$, $p<0.001$, $f=2.79$; $F_2(1, 187) = 254.039$, $p<0.001$, $f=1.17$] reflecting that word stimuli received faster responses than nonwords. All further analyses were performed for word and nonword targets individually.

There was no effect of congruency in word targets in *z*-scores [$F_1(1, 34) = 0.030$, $p=0.864$; $F_2(1, 96) = 0.002$, $p=0.964$]. This was also the case in nonword targets [$F_1(1, 34) = 0.219$, $p=0.643$; $F_2(1, 91) = 0.101$, $p=0.751$].

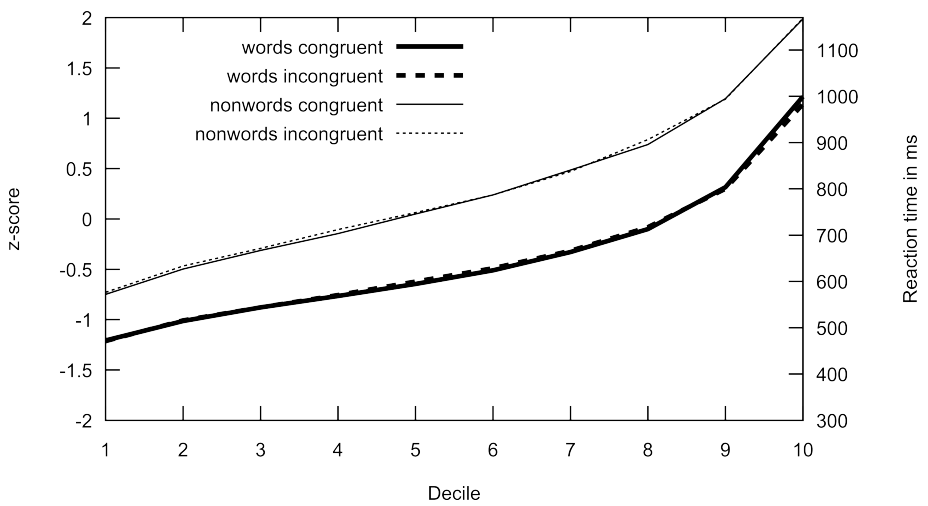

Figure 6.2: Plot of results of Experiment 7 by decile and condition. Note that the scale in ms is an approximation and data was plotted by z-scores.

The analysis of error ratio in word targets showed no effect of congruency [$F_1(1, 34) = 0.051$, $p=0.823$; $F_2(1, 96) = 0.053$, $p=0.818$]. But there was a significant effect of congruency on the error ratio in nonword targets [$F_1(1, 34) = 14.287$, $p=0.001$, $f=0.65$; $F_2(1, 91) = 12.979$, $p=0.001$, $f=0.38$], reflecting more erroneous responses to congruent trials than to incongruent trials.

Analysis in deciles

As in Experiment 1 the reaction times were binned in ten vincentiles and the results submitted to an analysis using lme4 in R (see Appendix E for a detailed list of results). The results showed that the congruency effect was significant in only one decile in word targets; specifically there was a significant negative effect in the slowest bin of responses. There was no significant effect in nonwords. This suggested that there was no detectable response congruency effect in this experiment, which was also compatible with the visual inspection of Figure 6.2 and the mean RTs in Table 6.3.

Ex-Gaussian

The data were analysed by fitting an Ex-Gaussian distribution to the RT of this experiment according to the method outlined in Experiment 1. The resulting parameters are shown in Table 6.4. There were no significant effects in word targets: μ [$F(1, 34) = 0.063$, $p=0.804$], σ [$F(1, 34) = 0.033$, $p=0.856$] and τ [$F(1, 34) = 0.336$, $p=0.566$]. The analysis in nonword targets also showed no significant effects: μ [$F(1, 34) = 0.116$, $p=0.735$], σ [$F(1, 34) = 1.713$, $p=0.199$] and τ [$F(1, 34) = 0.426$, $p=0.518$]. The absence of effects was in line with the analysis of variance and the analysis by deciles.

Table 6.4: Results of the Ex-Gaussian fitting of Experiment 7 using QMPE as a function of target lexicality and congruency.

	Word			Nonword		
Parameter	Incongruent	Congruent	*Effect*	Incongruent	Congruent	*Effect*
μ	500	501	*-1*	624	629	*-5*
σ	35	36	*-1*	64	82	*-18*
τ	138	134	*4*	182	172	*10*

6.3.3 Discussion

There was no congruency effect in this experiment in any of the analyses of RT. This was finally compatible with reports in the literature (Norris & Kinoshita, 2008; Perea et al., 1998, 2010). This experiment differed from Experiment 6 in both, nonword primes and nonword targets. The targets were replaced to further increase the difficulty level of the explicit task and the primes to reduce the prime informativeness. Only the combination of both measures produced a null effect whereas the five previous experiments showed a robust congruency effect. The next experiment will test whether this has to be attributed to the primes, the targets or both.

Replacing the nonword targets successfully increased the difficulty level of the task. This was illustrated by both an increase in RT and in error ratio compared to previous experiments. The error ratio in nonword targets was relatively high (6.75% in the congruent and 9.98% in the incongruent condition compared to 3.8% and 1.9% respectively in Experiment 2). There was also a strong effect in RTs (794 ms in nonwords and 637 ms in words compared to about 487 ms and 508 ms in Experiment 2) indicating the increased difficulty of the task. The RTs reported in this experiment were similar to those reported in the literature (see 1.4.2.3). As mentioned above, the nonwords in the study using English items (Norris & Kinoshita, 2008) were very wordlike and had a high similarity to the word targets (e.g., *STEAK* vs *STERK*). The items in the present experiment were not similar to each other on a pairwise level, but the nonwords were similarly wordlike on a whole. Thus, the word-nonword discrimination task was comparably difficult as indicated by slow RT and relatively high error ratios.

The absence of a prime congruency effect in the current results challenge an account that predicts that the priming effect increases in a more effortful task (Bodner & Masson, 1997), because response congruency effects were reduced with an increasing task difficulty and more effortful processing. The current results were compatible with both hypotheses that could explain reduced response congruency effects in more difficult tasks. The weight assigned to summed lexical activity could have shifted to identification as a result of the increased difficulty of the word-

nonword discrimination. Also, the absence of a response congruency effect could be reduced with a longer decision process that provides a longer time span for the effect to diminish or trickle out. Interestingly, this predicts that the congruency effect could be re-established in these targets with very informative primes. These primes would have a greater impact in early processing stages and thus, it takes longer for the effect to trickle out. In contrast, very informative primes could not re-establish the congruency effect, if the absence of an effect was due to a smaller weight assigned to summed lexical activity. The next experiment tested whether this is the case.

The current experiment used a very similar method of assigning prime and target pairs as Experiment 2. This means that stimuli were partly novel and partly used as targets prior to being employed as primes. The respective analysis of prime novelty (see Appendix A) did not reveal any reliable effects of used primes in Experiment 2. The current experiment did not show any effects using a similar method and strengthens the results of the analysis presented in Appendix A. The results suggested that using the primes as targets does not necessarily result in a stimulus-response mapping if there is a large number of targets (100 per category) and these are only presented once. However, stimulus-response mappings can occur in small sets (about a dozen targets) that are presented a few times (Damian, 2001). Importantly, the current experiment showed no sign of a stimulus-response mapping. Furthermore, it replicated the literature (Norris & Kinoshita, 2008; Perea et al., 1998, 2010) in terms of the magnitude of RTs and the absence of response congruency effects.

6.4 Experiment 8

This experiment tested whether a response congruency effect can emerge with targets that are relatively difficult when combined with very informative primes. Thus, in this experiment the same targets as in Experiment 7 and the same primes as in Experiment 2 were used. If a response congruency effect can be shown for the same set of targets used in Experiment 7, it is clear that prime informativeness is the critical factor.

6.4.1 Methods

Participants. Thirty-eight participants from the same population as in Experiment 1 took part in the experiment. All were native speakers of English.

Stimuli & Design. 100 word targets were taken from Experiment 4 and thus, the same as in Experiments 5, 6 and 7. The nonword targets were the same 100 wordlike stimuli used in Experiment 7. The prime stimuli were identical to those used in Experiment 2 through 5. In these experiments the same primes produced a stable congruency effect (see Appendix B for a list of stimuli). As in the prior experiments, primes and targets were paired so that they did not share any letters. In this experiment the prime and target stimuli were drawn from different sets. Two versions of the experiment were formed for counterbalancing purposes.

Procedure. The procedure was the same as in Experiment 2.

6.4.2 Results

Participants and items showing an error rate greater than 25% were dropped from the analysis. This affected two participants (43% and 27% error rate), four words (*motto*, *inert*, *bison*, *avail*) and five nonwords (*chave*, *becon*, *proad*, *quess*, *sping*). Of the remaining participants 18 had been assigned to each of the two lists. Outliers were removed by 3 SDs for each participant using the correct responses. This affected 1.73% of the remaining data. The mean RTs and error ratios are shown in Table 6.5.

Repeated measures analysis of variance

ANOVAs were performed using the participant-wise *z*-scores. In all analyses, the list factor was entered as a between participants factor and congruency was added as repeated within participants factors. Lexicality was a repeated within participants factor in F_1 and a between items factor in F_2. There was a main effect of lexical state in *z*-scores [$F_1(1, 34) = 228.575$, $p<0.001$, $f=2.60$; $F_2(1, 187) = 297.329$, $p<0.001$, $f=1.26$], reflecting that word stimuli received faster responses than nonwords in this experiment. All further analyses were performed for words and nonwords individually.

Table 6.5: Mean reaction times in ms and error ratio in percent as a function of lexicality and congruency of Experiment 8.

	Lexicality			
	Word		Nonword	
Congruency	RT	Error	RT	Error
Incongruent	601	4.40	713	7.72
Congruent	588	3.70	721	7.54
Effect	13	0.70	-8	0.18

The separate analyses of response congruency in word targets showed a significant main effect in *z*-scores [$F_1(1, 34) = 9.912$, $p=0.003$, $f=0.54$; $F_2(1, 94) = 6.860$, $p=0.010$, $f=0.27$] reflecting faster responses to congruent than to incongruent trials. In contrast to words there was no significant effect in nonword targets [$F_1(1, 34) = 2.140$, $p=0.153$; $F_2(1, 93) = 2.428$, $p=0.123$].

Analysis in deciles

The *z*-score transformed reaction times were binned in ten vincentiles and submitted to an analysis using lme4 in R (see Appendix E for a detailed list of results). The congruency effect in word targets was not significant in the first three deciles, but in all later deciles where congruent trials received faster responses than incongruent trials. This pattern was very similar to Experiment 6 where the congruency effect built up in the first half of the RT distribution. In nonword targets, there was no effect of

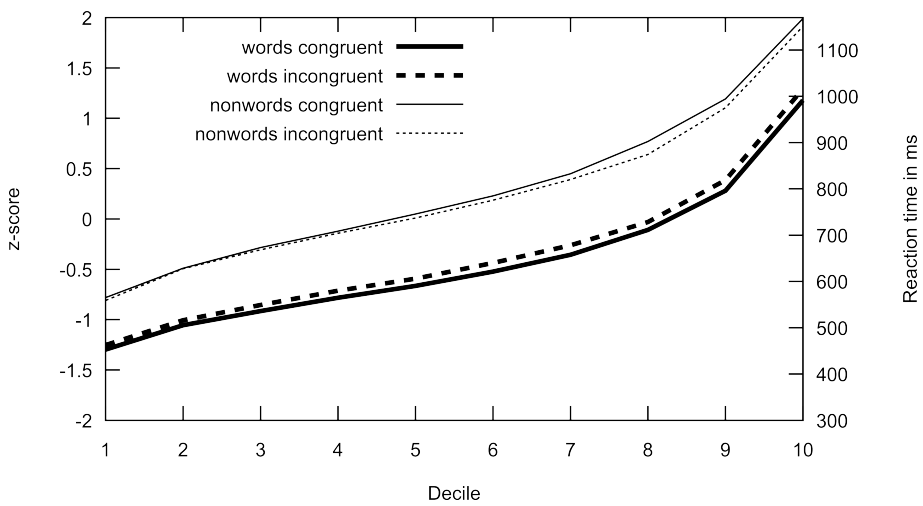

Figure 6.3: Plot of results of Experiment 8 by decile and condition. The scale of reaction times in ms is an approximation to reflect the actual corresponding value of z-score and reaction time.

congruency in the first seven deciles, but a negative effect in the last three deciles. That means slow responses to incongruent trials were faster than slow responses to congruent trials. This pattern was similar to the ones found in earlier experiments. In Experiment 3 and 6 there was a negative effect in slow responses and in Experiment 4 the effect became smaller.

Ex-Gaussian analysis

The data were analysed by fitting an Ex-Gaussian distribution to the RT of this experiment according to the method outlined in Experiment 2. The results are listed in Table 6.6. There was a significant main effect of congruency in word targets on μ [$F(1, 34) = 7.648$, $p=0.009$, $f=0.38$], reflecting faster responses in congruent trials. There was no effect in σ [$F(1, 34) = 0.557$, $p=0.460$] and τ [$F(1, 34) = 0.119$, $p=0.732$]. The analysis in nonword targets showed no significant effects in μ [$F(1, 34) = 0.099$, $p=0.755$] and σ [$F(1, 34) = 0.075$, $p=0.786$]. There was a tendency in τ [$F(1, 34) = 3.327$, $p=0.077$], reflecting more slow responses in congruent than in incongruent trials. This was compatible with the analysis in deciles, where a negative congruency effect emerged in the slower responses.

6.4.3 Discussion

This experiment showed a significant response congruency effect in word targets. Even though the explicit task was formed of the same items as in Experiment 7, a prime congruency effect emerged with an effect size similar to Experiment 5 and 6. The only difference between the current experiment and Experiment 7 was the prime informativeness. The more informative primes resulted in a congruency effect, but the less informative primes in Experiment 7 did not. This showed that the difficult targets

Table 6.6: Results of the Ex-Gaussian fitting of Experiment 8 using QMPE as a function of target lexicality and congruency.

	Word			Nonword		
Parameter	Incongruent	Congruent	*Effect*	Incongruent	Congruent	*Effect*
μ	496	484	*12*	585	583	*2*
σ	40	37	*3*	51	49	*2*
τ	105	103	*2*	134	150	*-16*

could be used in an experimental setting that produced a congruency effect. Hence, the distinctively important variable was the prime informativeness.

In nonword targets, there was no significant congruency effect in the repeated measures ANOVA, although there was a tendency in τ in the Ex-Gaussian fitting, indicating a longer tail in the RT distribution for congruent nonword trials than for incongruent trials. This finding was supported by a significant negative congruency effect in the slowest three deciles. The presence of these small effects in slow responses indicated that there was an effect of prime congruency.

With regards to the effect of prime repetition, this experiment used primes and targets drawn from disjoint sets, thus the primes never occurred as targets in this experiment. Despite a very difficult task, the novel primes resulted in a response congruency effect. This contrasted with Experiment 7 where the primes were partly used as targets prior to their usage as a prime, but no response congruency effect emerged. These findings strengthened the analysis in Appendix A and suggested that prime novelty showed no detectable effect in this experimental paradigm.

6.5 Discussion

The experiments in this chapter varied the informativeness of the primes with respect to the explicit task. In Experiment 6, the task difficulty was kept constant in comparison to Experiment 5, but the primes were less informative. The experiment revealed a smaller but significant response congruency effect in word targets. In Experiment 7, the nonword primes and targets were replaced by very wordlike items. This increased the task difficulty further, but it also reduced the prime informativeness. In this experiment, no response congruency effect was observed, replicating the null effects reported in previous experiments (Norris & Kinoshita, 2008; Perea et al., 1998, 2010). Also, the magnitude of the mean RT was similar to these experiments. Experiment 6 and Experiment 7 differed with respect to both prime informativeness and the difficulty of the explicit task. For this reason, Experiment 8 tested the same targets as Experiment 7, but used very informative primes. Under these conditions congruency effects were similar to those found in Experiment 6. This suggested that

congruency effects were influenced by the explicit task, specifically, the effect size decreased with increased task difficulty. The comparison of Experiment 7 and Experiment 8 suggested that the presence of a response congruency effect depended on prime informativeness. In Experiment 7 using less informative primes the manipulation of response congruency produced a null effect, but in Experiment 8 with very informative primes and the same targets, an effect similar to Experiment 6 was found.

6.5.1 Word targets

A similar congruency effect for word targets was observed in Experiment 6 and Experiment 8. In both experiments the congruency effect was not significant in the very fast responses. This is different from the experiments in the previous chapter where the effect showed a tendency to be stronger in fast than in slow responses. Furthermore, the congruency effect in the current two experiments did not decrease in size in slower responses.

The absence of congruency effects in fast *yes*-responses could reflect fast identifications of the target and thus, the effect of summed lexical activity was negligible compared to slower responses. This is compatible with Bodner and Masson (1997) who suggested that priming effects increase with more effortful processing. Specifically, the current data indicated that there was a lower boundary of processing effort before the primes could become effective. This minimum appeared to be relative to the respective explicit task, e.g. responses in Experiment 6 were generally faster than in Experiment 8, but the effect was building up in both experiments. Note that this is not incompatible with decreasing effects in slow responses. These were attributed to leakage in the decision channels or a reset mechanism reducing the observable effect in slow responses. Taken together, these findings implied that there is a window where congruency effects can be observed. The lower boundary appeared to be relative to the task requirements and the upper boundary was particularly influenced by the leakage in the decision system. On the other hand, the absence of response congruency effects in fast responses was only observed in two experiments

and it could also be the case that the congruency effects were not detected with the current methods.

6.5.2 Nonword targets

Using a standard analysis, the nonword targets did not show effects of response congruency in any of the three experiments reported in this chapter. It was only in a more fine-grained technique that negative congruency effects were revealed in slow responses in Experiment 6 and Experiment 8. Additionally, a very short-lived positive congruency effect was found in Experiment 6. Negative congruency effects in slow responses and positive congruency effects in fast responses were also reported in Experiment 3. A similar explanation can be applied across these three experiments. The nonword responses were amongst the slowest responses in all experiments in this chapter. In fast responses, only a small effect was measured because the primes were relatively uninformative, i.e. word primes resulted in a comparably little lexical activity and nonword primes produced a relatively large lexical activity. In slower responses, a negative effect emerged because the prime was processed up to a stage where it became a reasonable candidate for identification (inhibition threshold in Boy & Sumner, 2010), but never reached the identification threshold. This could have triggered a reset of the activation in the prime node and resulted in a negative congruency effect.

Interestingly, negative congruency effects in nonword targets only occurred in experiments where positive congruency effects were found in word targets. This was compatible with findings by Boy and Sumner (2010) showing that negative priming effects can only emerge if there is at least a potential for positive priming effects. It could be concluded that negative priming effects not only provide evidence for an effect of the priming conditions. They could also be interpreted as indicating that very similar processes as in positive priming effects were operating, but additionally inhibitory processes were active. For the current experiments it was suggested that there is a reset mechanism operating if a word node reached a required minimum activity and did not hit the identification threshold for a while.

Negative congruency effects occurred in nonword targets only. The hypothesis of a reset mechanism provides a simple explanation for this. The reset is only triggered after a certain amount of processing with some mismatch between the activated representation and the stimulus. The word targets received faster responses in all experiments in this chapter and thus, it could be argued they were clear cases and not subject to a reset. In other experimental paradigms where responses are equally difficult, negative priming effects occur with all response options (Boy & Sumner, 2010; Eimer & Schlaghecken, 2003; Kiesel et al., 2008).

6.5.3 Prime informativeness

The three experiments in this chapter tested the effect of prime informativeness on the presence of congruency effects. Two experiments (Experiment 6 and Experiment 8) using informative primes revealed very similar results and a similar effect size of response congruency. The third experiment used less informative primes and produced a null effect (Experiment 7). The evidence was in favour of the idea that response congruency effects depend on prime informativeness. Informativeness with regards to a lexical decision task was the amount of lexical activity that stems from the prime presentation, where a high level of activity was an indicator for a word response and a low level of activity indicated a nonword response. The important metric is the difference between the prime categories. If the nonwords were very wordlike and words low in frequency, both categories were almost indistinguishable with regards to the summed lexical activity resulting from the prime presentation. This situation was tested in Experiment 7, where response congruency effects were not obtained, supporting the idea that such effects only emerge if the word primes produce a higher level of summed lexical activity than the nonword primes like in Experiment 8.

This account of prime informativeness does not refer to the lexicality of the prime stimuli per se, but to the lexical activity that is triggered by their presentation. If this assumption is correct, response congruency effects could emerge independently of the actual lexical status of the prime stimulus. If typicality of nonwords is an approximation of the total lexical activity, then typical nonword primes would behave more similarly to word primes than to atypical nonword primes. Hence, using typical and atypical

nonword primes would produce similar effects like using word and nonword primes. The next chapter investigates this prediction.

6.5.4 Theoretical accounts of response congruency effects

The deep processing account (Dehaene et al., 1998; Naccache & Dehaene, 2001) predicts that the task instructions are applied to the prime. The processing of the prime stimulus is unconscious and incomplete, but essentially the same as the target processing. It results in motor readiness in one of the response options. In this account, the presence of a response congruency effect is independent of the difficulty of the explicit word-nonword discrimination task, but the activation in the respective response channels depends on the processing of the prime. Thus, the empirical findings in this chapter are compatible with this account. In Experiment 7 and Experiment 8, the explicit word-nonword discrimination was the same, but the very informative primes in Experiment 8 resulted in a response congruency effect whereas the less informative primes in Experiment 7 did not. This could be attributed to a less successful processing of the primes in Experiment 7. Indeed, if the primes of Experiment 7 are used as targets (e.g., Experiment 7 and 8) the response latencies are considerably longer than when the primes of Experiment 8 were used as targets (Experiment 2).

The negative congruency effects that were observed in the experiments in this and the previous chapter could be explained if the deep processing account was extended by a reset mechanism in the motor response channels. For example, if a word prime triggers lexical activity which in turn causes activity in the motor response channel, this would facilitate a word response and cause a conflict with a nonword response (positive effect). The timing of a reset is crucial in the negative effects as they only occurred in the slowest responses. Thus, the reset would operate after the average response time has passed, i.e. the slow nonword responses. At this time the word and the nonword response channel would have received considerate activation and a reset mechanism (e.g., a gated dipole field, Carpenter & Grossberg, 1987) would quickly allow the other channel to trigger a response, i.e. slow incongruent nonword trials

could benefit from activation in the word response channel compared to a slow congruent nonword trial.

The semantic overlap model (Quinn & Kinoshita, 2008) is compatible with the data, because it predicts that the congruency effect depends on the features that were activated. That means the word primes in Experiment 7 may not have activated the word-related features as strongly as the primes in Experiment 8 and vice versa the nonword primes in Experiment 7 could have activated some word-related features whereas the nonword primes in Experiment 8 did not. This hypothesis provides another interesting prediction. If a set of nonwords could be chosen in a way that they activate word-related features and another set that does not, the response congruency effects that were observed with words and nonwords should be replicated. Quinn and Kinoshita (2008) used impostor primes (nonexemplars that share central features with a category) and observed the predicted priming effects. This prediction was tested in the next chapter.

The results of Experiment 7 showed that even though the primes were drawn from the same pool of items as the targets there was no trace of a response congruency effect. That means that stimulus-response mappings do not build up from a single presentation of large set of targets, but after a few presentations (Damian, 2001). Also, the action trigger account cannot account for the data. If participants prepared themselves to respond to a certain set of features, then a stronger effect was expected in Experiment 7 where primes and targets were of a similar quality than in Experiment 8 where primes and targets differed with respect to their wordlikeness. But the empirical data showed that participants did not show a response congruency effect in Experiment 7 despite having the chance to adapt towards the targets which were also the primes. Rather a response congruency effect was observed in Experiment 8, where the set of targets that participants could adapt to was not similar to the set of primes.

6.6 Summary

The experiments in this chapter tested the impact of prime informativeness on response congruency effects. The informativeness of the prime towards a lexical

decision could be understood as difference in lexical activity that was triggered by the primes of in each condition. An estimate for this is the typicality or wordlikeness of a nonword (see Chapter 2.3 for a more detailed account). In a more typical lexical decision task than Experiment 2, response congruency effects were observed in Experiment 8 with informative primes, but not in Experiment 7 using less informative primes. This was the case even though both experiments used the same targets. Thus, prime informativeness determines whether a response congruency effect can occur in an experiment. The preliminary definition of prime informativeness refers to the lexical activity resulting from the prime, but not its actual lexical status. This suggests that a set of impostor nonwords (a set of very wordlike nonwords resembling words very closely) and a set of nonwords that is very unlike words could produce a wordlikeness congruency effect. The next chapter tests this hypothesis.

Negative congruency effects occurred only in nonword targets. These effects were only present when positive congruency effects emerged in word targets and could be taken as evidence for an effect of the priming conditions. The occurrence of negative congruency effects was attributed to a reset mechanism. The absence of negative effects in word targets was attributed to the generally faster responses so that they are not subject to this reset.

7. Effects of prime typicality

This chapter investigates prime typicality effects in masked primed lexical decision. The findings of the previous chapter showed that response congruency effects occurred when word and nonword primes were sufficiently informative. This manipulation confounded prime lexicality and prime typicality. In the previous experiments, the distinctive difference between these two priming conditions was to what extent the prime could bias the decision process in the one or the other direction. In particular, the lexical activity triggered by a very typical and high frequency word prime biased the decision process towards a word response. Nonwords that were very unlike words showed the opposite effect by biasing the process towards a nonword response. Woollams, Silani, Okada, Patterson, and Price (2011) showed in an fMRI study that a left inferior occipital region of the left ventral occipito-temporal cortex (LvOT) is particularly sensitive to typicality. Specifically, the effect of typicality was independent of the lexicality whose effect was stronger in the posterior LvOT. The following experiments investigated whether orthographic typicality was a component in word recognition that already contributed to lexical activity. Thus, it could be possible to form stimuli that do not differ in lexicality but the difference in typicality is sufficient to affect the response speed in an experiment. These stimuli would be similar to impostor primes (see 1.3.1.3).

Quinn and Kinoshita (2008) found that words that share central features with a category resulted in similar priming effects as actual members of the category. For example, Quinn and Kinoshita (2008) used body parts as a category. In their experiment the impostor *mind* resulted in similar priming effects as the actual exemplar *ear*, but the positive priming effect in exemplar targets was smaller with impostor primes than with exemplar primes. With respect to lexical decision, an impostor prime could be a nonword with features of a word or vice versa. Since most English words are typical for the English language to certain degree, it is easier to manipulate nonwords to resemble English words. It was expected that the results of using impostor nonwords compared to clear cut nonwords would be similar to using word and nonword primes respectively. Specifically, nonword primes that are very

typical for English are expected to trigger a high level of lexical activity compared to very atypical nonwords. That means that orthographic typicality allows an estimation of the chances of the target stimulus being a real word or unknown, independently of the lexicality of the prime. Thus, orthographic typicality (see Chapter 2) was the shared feature between actual category members (words) and the impostor primes (nonwords).

In the experiments reported so far, typicality and lexicality were confounded, because the word primes were always more wordlike than the nonword primes. Disentangling orthographic typicality and lexicality allowed a clear test of whether the congruency effects observed in the earlier experiments could be attributed to lexical activity that could also be triggered by a nonword prime. Furthermore, it provided a stringent test of the idea that any learnt associations between the stimuli and their responses could account for the response congruency effects reported in previous chapters. In the case of stimulus-response mappings all primes in the following experiments would have to be associated with a nonword response. Thus, a null effect would be expected if the congruency effects reported earlier could be attributed to response mappings. In contrast, a congruency effect had to be attributed to a stage earlier than the motor response and would support the hypothesis that orthographic typicality contributes to the word recognition process.

In the *low typicality* condition the nonword primes were very atypical compared to English words and hence were assumed to trigger a low level of lexical activity. In the *high typicality* condition the nonword primes resembled English words, that means they comprised typical letter combinations such as *th*, *ea*, *ight*. These nonwords formed the equivalent of impostors in a semantic categorisation task and were assumed to trigger a comparably high level of lexical activity. Finding a prime typicality effect could shed light on whether the congruency effect reported in previous chapters is due to congruence in responses or due to a bias from prime induced activity in the lexicon.

7.1 Experiment 9

The experiments in previous chapters suggested that response congruency priming effects are larger when the mean reaction time of responses is relatively fast. Also, within the RT distribution of a single experiment, the effect size typically decreased with slower responses. Thus, in order to maximise the odds of finding a typicality effect, the targets in this experiment were relatively easy to categorise. The word targets were of high frequency and the nonword targets were very unlike words, mirroring the target stimuli of Experiment 2. In order to disentangle lexicality and orthographic typicality, all primes were nonwords. In contrast to words, nonwords do not have lexical frequency allowing for the elimination of this confounding variable. Finding an effect of prime typicality would suggest that the response congruency effects reported above were due to the effect of the primes on the summed lexical activity as an indicator of lexicality.

In comparison to the Experiments 2 through to 8 that used word and nonword primes, the two categories of primes were more similar in this experiment. The reason was that a very wordlike nonword is expected not to activate the corresponding word node as much as the actual word. All lexical activity triggered by the nonword prime is the result of activity in word nodes that are similar to the stimulus. As a result it could be possible that a very wordlike nonword prime triggers a lower lexical activity than a word prime that was used in Experiment 2 and thus, the expected effect size was relatively small.

7.1.1 Methods

Participants. Thirty participants from the same population as in Experiment 1 took part in the experiment. All were native speakers of English.

Stimuli & Design. The word targets were taken from Experiment 2. These items were high in frequency and easy to categorise. The nonword targets were generated with similar restrictions as in Experiment 2. Their typicality was very low and they contained a combination of three letters that never occurred in English words (*OT3* could not be computed since the respective trigrams did not occur). Furthermore, it was ensured

that at least one vowel occurred in all items. The nonwords were also simulated using the Spatial Coding Model (SCM, Davis, 2010). Items with a σ value (reflecting lexical activity) of more than 0.1 after 50 cycles were excluded and replaced by other stimuli that fulfilled the restrictions. The primes were all nonwords. The primes with a low level of typicality were taken from Experiment 2. The primes with high level of orthographic typicality were taken from Experiment 7. Assuming that the difference between priming conditions in terms of lexical activity was crucial this choice maximised chances for finding an effect. The nonword primes in Experiment 2 were sufficiently different from word primes to contribute to a congruency effect. In contrast, the nonword primes in Experiment 7 were sufficiently similar to the word primes that no effect emerged. Thus, in this experiment one set of primes was very different from words and the other was as wordlike as possible. By generating another set of atypical nonwords for the use as targets, the primes and targets were recruited from different sets. Also, generating new targets allowed for using the primes that were shown to be effective in previous experiments without repetition (Appendix B for a list of all stimuli).

Primes and targets were paired so that there were no letters shared between prime and target. Two versions of the experiment were formed for counterbalancing purposes.

Procedure. The procedure was the same as in Experiment 2.

7.1.2 Results

As in all previous analyses, participants and items showing an error rate greater than 25% were dropped from the analysis. This did not affect any participants or items. Half of the participants were assigned to each of the two lists. Outliers were removed by 3 SDs for each participant using the correct responses. This affected 1.40% of the remaining data. The mean reaction times and the error ratios are presented in Table 7.1.

Repeated measures analysis of variance

ANOVAs were performed using the participant-wise *z*-scores. In all analyses the list factor was a between participants factor. Lexicality was a repeated factor in F_1 and

Table 7.1: Mean reaction times in ms and error ratio in percent as a function of lexicality and nonword prime typicality of Experiment 9.

	Lexicality			
	Word		Nonword	
Condition	RT	Error	RT	Error
Low typicality	517	4.33	544	1.20
High typicality	498	1.07	554	1.07
Effect	19	3.26	-10	0.13

between items in F_2 and prime typicality was added as a repeated within participants factor. There was a main effect of lexical status in *z*-scores [$F_1(1, 28) = 33.734$, $p<0.001$, $f=1.10$; $F_2(1, 196) = 78.746$, $p<0.001$, $f=0.63$] reflecting that word stimuli received faster responses than nonwords in this experiment. All further analyses were performed for word and nonword targets individually.

In word targets the prime typicality effect was significant in *z*-scores [$F_1(1, 28) = 27.351$, $p<0.001$, $f=0.99$; $F_2(1, 98) = 47.387$, $p<0.001$, $f=0.70$] indicating that word targets received faster responses when they were preceded by a high typicality than by a low typicality prime. The prime typicality effect was significant in nonword targets as well [$F_1(1, 28) = 13.770$, $p=0.001$, $f=0.70$; $F_2(1, 98) = 32.747$, $p<0.001$, $f=0.58$]

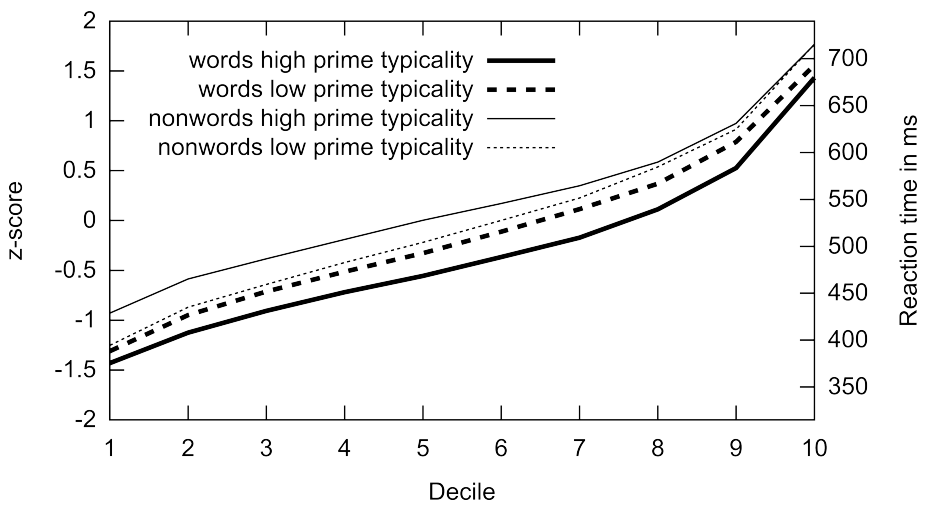

Figure 7.1: Plot of the results of Experiment 9 by decile and condition. The scale of reaction time in ms is an approximation to reflect the corresponding values.

showing that nonwords were responded faster when they were primed by low typicality stimulus than by a high typicality stimulus.

The analysis of error ratio in word targets showed a main effect of prime typicality [$F_1(1, 28) = 64.149$, $p<0.000$, $f=1.51$; $F_2(1, 98) = 23.118$, $p<0.000$, $f=0.49$] indicating more errors in the low typicality priming condition than in the high typicality condition. There was no such an effect in nonword targets [$F_1(1, 28) = 0.111$, $p=0.742$; $F_2(1, 98) = 0.181$, $p=0.671$].

Analysis in deciles

The *z*-score transformed reaction times were binned in ten vincentiles before being submitted to an analysis using lme4 in R (see Appendix E for a detailed list of results). The effect of prime typicality was significant in all vincentiles in word targets indicating faster responses to word targets with a high typicality nonword prime than to words with a low typicality prime. In nonword targets the effect was significant in the faster seven deciles indicating faster responses to nonwords primed with a low typicality nonword than to nonwords primed with a high typicality nonword. The effect faded out in slower nonword responses.

Ex-Gaussian analysis

The RT data were fitted to an Ex-Gaussian distribution as outlined in Experiment 1 and the resulting parameters are shown in Table 7.2. There was a main effect of prime typicality in word targets on μ [$F(1, 28) = 13.624$, $p=0.001$, $f=0.70$] reflecting faster responses to high typicality primed words than to low typicality primed words. There was a strong tendency on σ [$F(1, 28) = 4.079$, $p=0.053$, $f=0.38$] indicating a greater

Table 7.2: Results of the Ex-Gaussian fitting using QMPE as a function of target lexicality and prime typicality of Experiment 9.

	Word			Nonword		
Parameter	Low typicality	High typicality	*Effect*	Low typicality	High typicality	*Effect*
μ	449	423	*26*	457	479	*-23*
σ	52	42	*10*	51	41	*10*
τ	68	77	*-9*	94	77	*17*

variance in the low typicality condition and no effect on τ [$F(1, 28) = 1.121$, $p=0.299$]. Similarly, a main effect of prime typicality was found in nonword targets on μ [$F(1, 28) = 15.877$, $p<0.001$, $f=0.75$] showing faster responses to nonwords with a low typicality prime than with a high typicality prime. There was also an effect on σ [$F(1, 28) = 6.427$, $p=0.017$, $f=0.48$] indicating a greater variance in response times to low typicality primed nonwords and no effect on τ [$F(1, 28) = 2.151$, $p=0.154$].

7.1.3 Discussion

The results of this experiment showed a typicality-lexicality interaction in word and nonword targets. Word responses were facilitated by a high typicality nonword prime compared to a low typicality prime, whereas nonword responses were facilitated by low typicality primes compared to high typicality primes. Thus, the results showed that high typicality nonwords worked as impostors by facilitating word responses and inhibiting nonword responses. High typicality nonword primes triggered a higher level of lexical activity than low typicality primes. A higher lexical activity contributed stronger to the *yes*-channel and hence, word responses were facilitated and nonword responses inhibited. This is compatible with Woollams et al.'s (2011) finding that typicality and lexicality formed separable properties of a letter string, but it also showed that typicality contributed to recognise a letter string as a word. In the previous experiments, it was assumed that the response congruency effect was triggered by a higher level of lexical activity induced by the prime. This hypothesis was supported by the current results. The task in the current experiment was easy and participants could rely on a global measure of lexical activity rather than identifying the target word (e.g., Grainger & Jacobs, 1996). Under these conditions the impact of the priming conditions mediated through global lexical activity was maximised and this facilitated the observed impostor effect.

With regards to the central tendency of the RT distribution the effect in ms in the present experiment was as large as in Experiment 2. Also, the magnitude of the typicality effect was similar in word and nonword targets resembling another result of Experiment 2. Despite a strong similarity of the RT distribution of word and nonword targets in the central tendency, the mean values were less similar. In nonwords, the

effect faded out towards the end of the distribution, but it remained stable in word targets. This was illustrated by not significant effects in the slowest three deciles in nonword responses and also by Figure 7.1. The decline of the effect in slow responses reduced the mean typicality effect to 10 ms compared to 23 ms in the central tendency. A reduced priming effect in slower responses in nonwords was observed in Experiment 3, 4, 6 and 8 as well. These findings were consistent with the suggestion of a reset mechanism that is activated only after some time has passed. Due to the time course of the reset mechanism the word responses that were generally faster were less or not affected. Thus, the response congruency effect was reduced in nonword targets but not in word targets.

This experiment showed that a congruency effect can not only emerge as a result of congruence in the actual responses, but also as a result of a shared level of activity. That means primes associated with a low level of lexical activity facilitated responses to targets with a low level of activity, i.e. nonwords. In contrast, primes triggering a high level of lexical activity facilitated responses to word targets. This pattern forms an interaction of target lexicality and the orthographic typicality of the prime. The target stimuli in this experiment were easy to categorise and it could be argued that participants were not identifying the targets but rather judged the orthographic legality of the strings. For example, Yap et al. (2006) showed how nonword legality modulates the lexical frequency effect in word targets. Thus, it could be argued that this experiment differed from a standard lexical decision task and that the categories in this experiment were actually less complex than assumed, i.e. participants were judging legality rather than lexicality. The next experiment tested this question by using targets that were more difficult to categorise which resembles a standard lexical decision task.

7.2 Experiment 10

Experiment 9 showed that congruency effects in masked priming can emerge as a result of matching or conflicting levels of typicality. The findings suggested that the typicality of a letter string forms an estimate of the lexical activity triggered by its presentation. The target stimuli in the previous experiment were very easy to categorise and participants could have responded on the basis of quickly judging the

orthographic legality of the presented stimuli. If this was the case the observed typicality-lexicality interaction would not necessarily involve lexical activity, but could be due to some quick unconscious judgement of whether the stimulus is language like or not. Under these assumptions the crucial feature was not lexical activity but the legality shared among words and high typicality nonword primes. The following experiment used a more difficult explicit word-nonword discrimination task to hinder participants from relying on legality and tested this potential rejection. The word targets were taken from Experiment 4 and the nonword targets from Experiment 3. This formed an experiment where the explicit task required a more sound word processing. An interaction of orthographic typicality and target lexicality in this experiment would support the account of lexical activity patterns. The effect size in this experiment was expected to be smaller than in Experiment 9, because participants had to rely on lexical identification to a greater extent. Thus, the impact of summed lexical activity in the decision process could be reduced and as a result the effect size was expected to be small. Most importantly, the legality account predicted a null effect in this experiment and finding an effect would severely challenge this idea.

7.2.1 Methods

Participants. Thirty-eight participants from the same population as in Experiment 1 took part in the experiment. All were native speakers of English.

Stimuli & Design. The word targets from Experiment 4 and the nonword targets of Experiment 3 were used. These items were harder to categorise than the targets in Experiment 2 and Experiment 9 respectively. These items were selected to ensure that participants could not rely on judging the legality of the stimuli and hence, avoided a possible rejection to the conclusions. The primes were taken from Experiment 9. Hence, the primes in the low typicality condition were the same nonword primes as in Experiment 2 and the primes in the high typicality condition were the same as in Experiment 7.

Primes and targets were paired so that there were no letters shared between prime and target (see Appendix B for a list of stimuli). Two versions of the experiment were formed for counterbalancing purposes.

Procedure. The procedure was the same as in Experiment 2.

7.2.2 Results

Participants and items showing an error rate greater than 25% were dropped from the analysis. This criterion did not affect any participants, but three word items (*avail*, *bison*, *inert*). Half the participants were assigned to each of the two lists. Outliers were removed by 3 SDs for each participant using the correct responses. The mean reaction times and the error ratios are presented in Table 7.3.

Repeated measures analysis of variance

ANOVAs were performed using the participant-wise *z*-scores. In all analyses the list factor entered as a between participants factor. Lexicality was a repeated within participants factor F_1 and a between items factor in F_2. Prime typicality was a repeated within participants factor. There was a main effect of lexicality in *z*-scores [$F_1(1, 36) = 20.765$, $p<0.001$, $f=0.76$; $F_2(1, 193) = 11.982$, $p=0.001$, $f=0.40$] reflecting that word stimuli received faster responses than nonwords. All further analyses were performed for word and nonword targets individually.

There was no significant effect of prime typicality in word targets in *z*-scores [$F_1(1, 36) = 0.879$, $p=0.355$; $F_2(1, 95) = 0.520$, $p=0.473$]. The prime typicality effect was not significant in nonword targets as well [$F_1(1, 36) = 0.739$, $p=0.396$; $F_2(1, 98) = 0.618$, $p=0.434$].

Table 7.3: Mean reaction times in ms and error ratio in percent as a function of lexicality and nonword prime typicality of Experiment 10.

	Lexicality			
	Word		Nonword	
Condition	RT	Error	RT	Error
Low typicality	567	5.59	598	2.89
High typicality	564	4.49	602	2.32
Effect	3	1.1	-4	-0.57

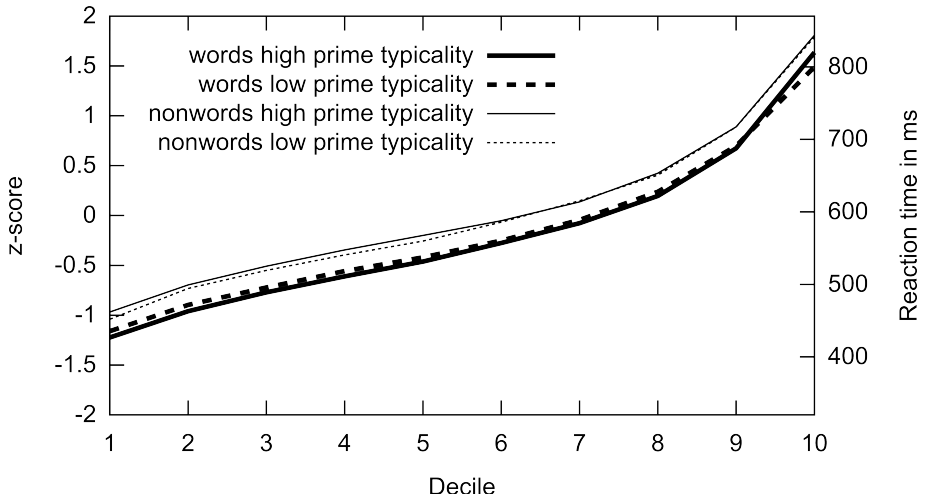

Figure 7.2: Plot of the results of Experiment 10 by decile and condition. The scale of reaction time in ms is an approximation to reflect the corresponding values.

The analysis of error ratio revealed no significant effect of prime typicality in word targets [$F_1(1, 36) = 0.793$, $p=0.379$; $F_2(1, 95) = 0.499$, $p=0.482$] and in nonword targets [$F_1(1, 36) = 1.840$, $p=0.183$; $F_2(1, 98) = 1.216$, $p=0.273$].

Analysis in deciles

The *z*-score transformed reaction times were binned in ten vincentiles before being submitted to an analysis using lme4 in R (see Appendix E for detailed list of results). The effect of prime typicality was significant in word targets in the fastest two vincentiles indicating faster responses to word targets with a high typicality nonword prime than to words with a low typicality prime. There was a significant negative effect in the slowest bin of responses to word targets. In all other deciles the difference between the conditions was numerically in the positive direction. The small negative effect indicated that in the Ex-Gaussian analysis, an effect in τ could occur and this could have covered the effect in mean RT. In nonword targets the effect was significant in the fastest decile and in decile four where low typicality primes facilitated nonword responses compared to high typicality primes. Finding a small effect in the fast responses was also compatible the with visual inspection of Figure 7.2.

Ex-Gaussian analysis

The RT data of correct responses were fitted to an Ex-Gaussian distribution using QMPE. The results are shown in Table 7.4. There was a main effect of prime typicality in word targets on μ [$F(1, 36) = 5.475$, $p=0.025$, $f=0.39$] indicating faster responses to words preceded by high typicality primes compared to words preceded by a low typicality prime. There was no effect on σ [$F(1, 36) = 0.939$, $p=0.339$] and no effect on τ [$F(1, 36) = 0.798$, $p=0.378$]. There were no significant effects in nonword targets: μ [$F(1, 36) = 2.082$, $p=0.158$], σ [$F(1, 36) = 0.025$, $p=0.874$] and τ [$F(1, 36) = 0.641$, $p=0.429$]. The effect on μ was not significant, despite a numerical difference showing an advantage of nonword responses in the low typicality condition over the high typicality condition that was almost as large as in word targets.

7.2.3 Discussion

The analysis of this experiment revealed a small effect of prime typicality in word targets. The effect in words diminished towards the end of the RT distribution. Analysing the effect in deciles showed that the effect was negative in the slowest bin of responses. Importantly, the analysis using the Ex-Gaussian fitting procedure revealed a significant positive effect of prime typicality in the central tendency of the RT distribution. Using this method the number of very slow responses was reflected in τ and allowed a clearer picture of the more central part of the RT distribution. In nonword targets there was a numerical trend in mean RT and the central tendency of the Ex-Gaussian fitting. Even though this trend was numerically similar to word targets, it was not significant in nonword targets.

Table 7.4: Results of the Ex-Gaussian fitting using QMPE as a function of target lexicality and prime typicality of Experiment 10.

	Word			Nonword		
Parameter	Low typicality	High typicality	*Effect*	Low typicality	High typicality	*Effect*
μ	471	461	*10*	493	502	*-9*
σ	39	34	*5*	37	38	*-1*
τ	98	103	*-5*	109	103	*6*

The small effect in word targets showed both that varying the primes' orthographic typicality was effective in a more difficult lexical decision task than Experiment 9, but it also demonstrated the impact of increasing the task difficulty. In a more difficult task the participants had to rely on identification to a greater extent than on global lexical activity (Grainger & Jacobs, 1996). This contrasts with easier tasks such as Experiment 2 and Experiment 9. Thus, the impact of priming conditions that differ on the dimension of global lexical activity was limited. This explained the smaller effect in the current experiment compared to Experiment 9. Importantly, the effect in this experiment could not be attributed to participants judging the legality of the presented letter strings and this prevented the findings of this experiment from a potential rejection to Experiment 9.

The small effect size in this experiment could be directly related to the time scale of lexical identification. As soon as a word target has been identified, a strong input feeds into the *yes*-channel. From this point in time the impact of global lexical activity is negligibly small and hence, the effect of the prime heavily reduced, especially in slow responses. The result of such an interaction of mechanisms is the observed pattern of a typicality priming effect that is strongly reduced towards the end of the RT distribution in word targets.

7.3 Discussion

Both experiments in this chapter demonstrated a typicality-lexicality interaction in lexical decision where an orthographically typical prime facilitated word responses and an orthographically atypical prime facilitated nonword responses. This evidence was in favour of the hypothesis that not the lexicality of primes but the activity they triggered in the lexicon was responsible for the response congruency effects observed in Experiment 2 to 8. Furthermore, this typicality-lexicality interaction cannot be attributed to stimulus-response mappings (see Damian, 2001), because all primes in Experiment 9 and 10 were nonwords and their effect would not result in an observable difference between the conditions. This supported the argument that the congruency effects reported in this thesis were not due to any learnt responses. Thus, both experiments provided evidence for an impostor effect (see Quinn & Kinoshita, 2008) in

lexical decision where an impostor prime refers to a nonword that was sufficiently wordlike to trigger effects that were more similar to a word prime. Also, this finding was compatible with Woollams et al.'s (2011) results in showing that the effects of lexicality and typicality can be disentangled. Additionally, the experiment showed that typicality contributed to the lexical decision process.

This typicality-lexicality interaction showed an effect size in Experiment 9 that was comparable to Experiment 2 and hence, one of the largest effect sizes amongst the experiments reported in this thesis. In Experiment 10 the effect size was considerably smaller and this was directly attributed to the difficulty of the task as this was the only difference between the two experiments. This result also supported the findings of Chapter 5 where an increasingly difficult word-nonword discrimination was associated with a decreasing effect size in response congruency effects. The difficulty of the explicit task was related to the weight that is assigned to summed lexical activation by the participants (Grainger & Jacobs, 1996). In priming effects that are mediated through the summed lexical activity as in all experiments introduced so far, a reduced weight in the decision process results in a smaller effect. This was compatible with the empirical findings of the experiments in this chapter.

Negative typicality effects

In Experiment 10, negative priming effects were observed in word targets whereas in previous experiments only nonword targets were associated with negative effects, e.g. in Experiment 3. But the underlying mechanism for the reversal in the direction could be very similar in word and nonword targets. Indeed, if a highly typical nonword prime triggered some lexical activity and in turn this triggered a reset, then this typical prime would facilitate a nonword response and interfere with a word response. This mechanism could also explain the fading out of effects towards the slower end of the RT distribution, because activity in a word that is due to a similar nonword stimulus is likely to grow more slowly compared to the actual word stimulus. Thus, a reset due to a mismatch is more likely in slow responses.

According to Boy and Sumner (2010) the presence of a negative effect also points on a potential positive effect, because self-inhibition can only occur if there was activation in the channel (for similar argument see Eimer & Schlaghecken, 2003; Schlaghecken et al., 2006). In all experiments reported in this thesis a negative effect was accompanied by a positive effect which strengthens this point. Thus, the negative effects also supplied evidence in favour of the response congruency and the impostor effect in masked primed lexical decision.

Theoretical accounts of response congruency

The experiments in this chapter used primes that were comparable to impostor primes (Quinn & Kinoshita, 2008) and thus, the semantic overlap hypothesis could provide an account for these findings. In this account the words can be described by relevant features. If the prime activated some of these features, a word response is facilitated and a nonword response is slowed down. Importantly, these word features must overlap between the prime and the target, but there was no semantic relation between primes and targets and no meaning associated to the nonword primes in the experiments. Thus, the word features appear to be features about the letter string, e.g. *pronounceable*, *legal*. In capturing orthographic typicality these features could relate to highly frequent letter combinations, but in the experiments in this chapter and all previous chapters, the primes and targets did not share any letters. Typicality could be represented by a single feature where its activation reflects the degree of typicality in the letter string which would enable the overlap hypothesis to account for the findings. In summary, the semantic feature overlap hypothesis provided an interesting view, specifically with regards to the impostor priming, but the difficulty is defining appropriate features to distinguish between words and nonwords that would still accommodate all findings in the empirical data.

Another account, that was able to accommodate the findings in the previous chapters is the deep processing account (Dehaene et al., 1998). In this account it is assumed that the task instructions are applied to the prime in the same as to the target, but the processing of the prime is incomplete. The results of the experiments in this chapter showed that primes that are associated with a *no*-response can facilitate a

yes-response. Although, the deep processing account assumes that the corresponding responses are activated and thus, all nonword primes should have pre-activated the *no*-response. It could be argued that the nonword impostor primes were indeed very wordlike and easily confused with words. This claim would be supported by the findings of Experiment 7 where a similar set of nonwords resulted in a comparably high error score and also findings by Quinn and Kinoshita (2008) who showed that in a speeded classification task impostors were very error prone but the accuracy was very high in unspeeded task. Thus, the recognition system would indeed apply the task instructions to all primes, but in the early stages of processing impostor primes fooled the system into activating the word response, which was the underlying reason for the difficulty of classifying impostors correctly. With this assumption the deep processing account would be able to explain the findings of Experiment 9 and 10.

7.4 Conclusion

The experiments in this chapter demonstrated an impostor effect in lexical decision, that means highly typical nonword primes facilitated a word response compared to less typical primes and vice versa low typicality nonword primes facilitated a nonword response compared to high typicality primes. In the previous experiments, a response congruency effect emerged as a result of manipulating prime lexicality and the typicality of the primes, but Experiment 9 and 10 provided evidence that these effects were triggered by the orthographic typicality of the primes independently of their lexicality. This suggested that the effect was not due to a congruence in responses between prime and target, but rather reflected the impact of the prime on a global measure of lexical activity. Thus, high typicality nonword primes can function like impostors in a lexical decision task, where they facilitated word responses.

In Experiment 9, the effect of prime typicality was comparable to Experiment 2 where the largest response congruency effect was reported in this thesis. Experiment 10 provided further support for the assumption that the response congruency effects were the result of prime induced lexical activity, because the increased difficulty of the explicit task may have resulted in a lower weight assigned to summed lexical activity (see e.g., Grainger & Jacobs, 1996). Thus, the effect size of prime typicality was

expected to be reduced, but not absent. The results supported this expectation. This also supported the conclusion from Chapter 5, where a smaller effect was reported with increasing task difficulty which was attributed to a shift in weight assigned to summed lexical activity. Importantly, the impostor effect in this chapter and the response congruency effect in Chapter 5 were still significant in a more difficult task. Thus, the observed effects cannot be explained by participants relying their lexical decision on legality rather than lexicality, because the targets in the explicit task did not allow a judgement based on legality.

A negative priming effect occurred in very slow responses to word targets in Experiment 10. The presence of negative priming effects implied the potential for positive priming effects (Boy & Sumner, 2010; Eimer & Schlaghecken, 2003; Schlaghecken et al., 2006). Thus, the negative priming effects strengthened the conclusion drawn from Experiment 10 that response congruency effects are induced by lexical activity triggered by the prime. Also, this provided further evidence for a reset mechanism as outlined in Chapter 5.5.2.

8. Simulations in the Spatial Coding Model

This chapter introduces a method of fitting the Spatial Coding Model (SCM; Davis, 2010) to empirical data. The results of the simulations of Experiment 2 through to 10 and the parameters from the model fits are discussed in this chapter.

8.1 Verification in the Spatial Coding Model

The empirical data reported so far showed that a manipulation of response congruency and prime typicality can result in positive, null and negative effects. A positive effect refers to a situation where a highly typical prime (or a word prime) facilitates word responses compared to a low typicality prime (or a nonword prime), or a low typicality prime facilitates *no*-responses compared to a high typicality prime. A negative effect refers to the opposite situation where a high typicality prime facilitates *no*-responses compared to a low typicality prime. The results of Experiment 2 to 6 and Experiment 8 showed positive response congruency effects in mean RT where prime typicality and lexicality coincided, i.e. word primes were typical and nonword primes atypical. There was a positive impostor effect (typicality-lexicality interaction) in Experiment 9 and 10. But no effect was observed in Experiment 7.

A more fine grained analysis of the data revealed a sharp decline of the response congruency effect in the slower responses of Experiment 3. The fine grained analysis of the RT distribution in deciles showed that there was a negative effect of typicality in the slowest two deciles of the RT distribution (see Figure 5.1). In other experiments, small negative congruency effects were observed in slow responses as well, this includes the nonword targets in Experiment 4, 6 and 8. In Experiment 10, the typicality effect in word targets turned into a negative effect in slow responses as well. Furthermore, there were cases where the effect strongly declined towards the slower end of the RT distribution, but did not turn into a negative effect; this pattern was observed for the nonword targets in Experiment 9 and the word targets in Experiment 4 and 5. Because negative priming effects and a sharp decline of the typicality effect occurred in a number of experiments with both word and nonword targets, a

satisfactory computational model of masked priming should be able to explain these findings.

One way to capture negative priming effects is by assuming a reset mechanism (see Chapter 5.5.2). The presentation of a highly typical prime (or a word prime) is assumed to activate word nodes that were similar to the prime stimulus, and this provides the source of the increased level of summed lexical activation that is assumed to trigger positive typicality effects. According to the reset account, once the *yes*-decision threshold is reached or a word node has hit the identification threshold, a verification process checks the correspondence between the word coded by the most active node and the pattern at the letter level. In the present experiments, the primes were never similar to the targets. If lexical activity triggered by the prime and activity by a nonword target were sufficient to trigger a *yes*-response on the basis of summed lexical activity, this verification process should always result in a mismatch and in turn, a reset of the most active word node. This verification-reset cycle is essentially the same as that proposed in adaptive resonance theory (e.g., Carpenter & Grossberg, 1987).

Resetting the most active word node results in an abrupt change of input into the decision channels, which could explain the effects that were observed in word and nonword targets. A highly typical prime contributes to summed lexical activity, whereas an atypical prime does not. In case the target is a nonword, the contribution of a highly typical delays the lexical decision. This is a situation where a positive priming effect is expected. Only if the lexical activity resulting from the prime and the target is sufficiently high a verification is triggered. That means that the verification process is triggered faster with a highly typical prime than with an atypical prime. As a result the verification mechanism turns an expected positive effect into a negative priming effect and thus, a negative effect can only occur if there was a potential for a positive effect. The process is similar in word targets, if there is another similar word that is higher in frequency. The verification mechanism could reset this related word node and delay a correct *yes*-response on the basis of a high level of summed lexical activity. In sum, the verification mechanism can produce negative effects in a situation where a positive effect was expected otherwise. This is compatible with the conditions on self-inhibition

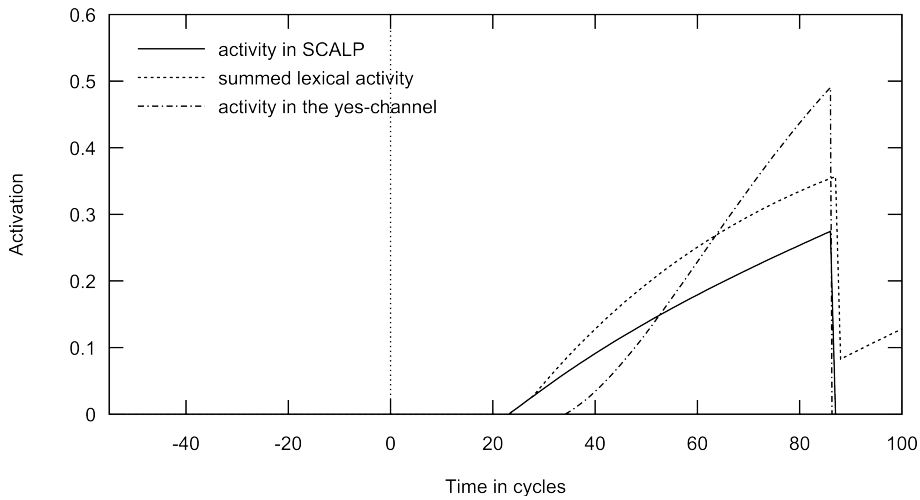

Figure 8.1: The flow of activity in the low typicality priming condition of Experiment 10. The data was derived using the example dqrki – SAULP.

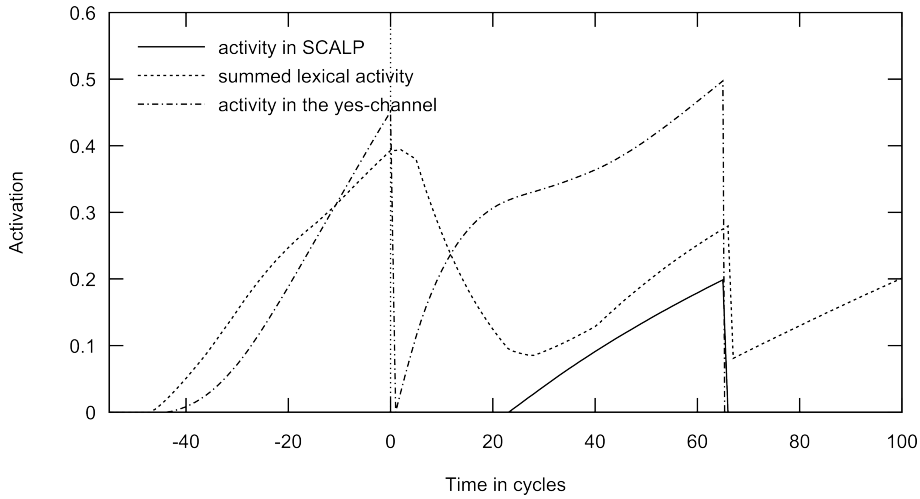

Figure 8.2: The flow of activity in the high typicality priming condition of Experiment 10. The data was derived using the example goven – SAULP.

formulated by Boy and Sumner (2010). In their account negative effects are due to an additional self-inhibition mechanism and can only occur if there was a potential for a positive effect. In the empirical data presented so far, positive and negative effects occurred within the same RT distribution and none of the experiments revealed a negative effect without showing a significant positive effect.

The relation of negative and positive priming effects is exemplified in Figure 8.1 and 8.2. In the low typicality priming condition (Figure 8.1), there is almost no lexical activity during the processing of the prime (cycles with negative numbers). The nonword target *SAULP* is similar to *SCALP*. Thus, the activity in the word node *SCALP* and lexical activity increases after target onset. This continues until a verification is

triggered in about cycle 90. In contrast, the verification is triggered in about cycle 65 with a high typicality prime (Figure 8.2). This is due to prime induced lexical activity which contributed to hitting the verification threshold. As a result, the reset of the word node *SCALP* is performed earlier and the correct *no*-response is triggered faster with the high typicality prime *GOVEN* than with the low typicality prime *DQRKI*. The high typicality prime resulted in a higher level of summed lexical activity. The positive priming effects that were observed in Experiment 2 were attributed to the higher level of summed lexical activity. But using a wordlike target such as *SAULP*, the verification threshold was hit and a reset of the most active word node performed. Thus, the additional verification process produced a negative priming effect.

In computational modelling of word recognition, the activation-verification model (Paap et al., 1982) suggested that an entry that was initially activated is verified in a second pass. The verification mechanism checks whether the activated entry matches the stimulus and, if required, resets the entry. This mechanism enables the model to correctly detect a mismatch between a stimulus and a very similar representation, i.e. correctly rejecting a nonword (false pretenders) that is very similar to an actual word. This verification mechanism was originally not intended to enable the model to predict negative priming effects, but rather to hinder the erroneous recognition of a word with a false pretender. Grainger and Jacobs (1999) also argued that a reset mechanism is important in models of word recognition. They pointed out that in normal reading words are identified in fast succession and the model needs a way of deactivating the word node corresponding to the previous stimulus. Grainger and Jacobs (1999) argued that the mismatch between the stimulus and the previously activated word node should trigger a reset and this would clear the activity of the preceding stimulus. This suggestion is compatible with the adaptive resonance theory (Carpenter & Grossberg, 1987) but also with the suggestion above.

In summary, the verification mechanism described above could enable the SCM to predict negative effects in simulating the present data, but also to deal with false pretenders. Thus, the verification was implemented in the model. Due to changes in the model's internal equations, the parameter specifying the inhibition between the

decision channels (λ) was set to 0.110 as a result of scaling the values accordingly. This change in the parameter settings will not be reported in the following simulations as it served a scaling purpose after incorporating the verification into the model and thus, is not of theoretical interest. The verification mechanism was activated in all following simulations with SCM.

8.2 Fitting the Spatial Coding Model

The Spatial Coding Model (SCM) provided very good correlation with empirical data in unprimed and masked primed lexical decision experiments (Davis, 2010). The current investigation did not rely on a single set of parameters, but the SCM was fitted to the empirical data by varying the parameters of the model. The reason for varying the parameters systematically was that the experiments varied considerably in task difficulty and it was hypothesised that participants shifted the weights assigned to the summed lexical activity and word identification accordingly. Thus, for testing this hypothesis the model's parameters were fitted to capture the empirical data. This section describes how the SCM was fitted to the empirical data.

8.2.1 Empirical data

The current implementation of the SCM does not produce an RT distribution for an item and does not try to explain a distribution of RT like the diffusion model (Ratcliff, 1978; Ratcliff et al., 2004). Rather, the response times of the model are deterministic. Thus, the simulations attempted to model the central tendency μ in the data rather than the entire RT distribution. The Ex-Gaussian (Heathcote et al., 2004) analysis of the empirical data showed that the central tendency of the RT distribution μ was less biased by the tail of the distribution than was mean RT and also very informative about the effects, particularly in Experiment 10. In this experiment an impostor effect was detected in μ, whereas this effect was disguised by the slower part of the distribution in the mean RT (also see Yap, Balota, Tse, & Besner, 2008). These data were derived for each experiment using QMPE (S. Brown & Heathcote, 2003) and used for fitting the model. All fits reported in the next section were derived by comparing the model output to the μ parameters estimated from the Ex-Gaussian analysis rather than the mean RT.

8.2.2 Parameters

The parameter set used in C. J. Davis (2010) formed the basis for the following simulations (see Appendix D). For the simulations and deriving the model fit, three parameters were systematically varied: the weight assigned to summed lexical activity (y_{global}) and the thresholds of the decision channels (Θ_{yes}, Θ_{no}).

The lexical decision experiments systematically differed in the level of difficulty in the explicit task. Grainger and Jacobs (1996) argued that the reliance on summed lexical activity accounted for differences in their experimental results when they manipulated the nonword foils between experiments ranging from wordlike to unwordlike items. In the SCM, the y_{global} parameter allows manipulating the degree to which summed lexical activity influences the decision channels. If this parameter is high then the wordlikeness of the target has a great impact on the decision and a decision could be triggered prior to the identification of the target. In an experiment with very wordlike nonword foils, this mechanism could result in false *yes*-responses. Thus, for simulating experiments where the word and nonword targets were very similar with respect to their wordlikeness, the y_{global} parameter should not be too high. This means that the model would trigger *yes*-decisions predominantly when the target word was identified and thus, avoid premature *yes*-responses to nonword targets.

In addition to y_{global}, the decision thresholds (Θ_{yes}, Θ_{no}) were fitted to the empirical data. Whilst the weight of summed lexical activity manipulates the sort of information feeding into the decision channels, the decision threshold determines the amount of evidence required for triggering a response. The experiments presented in this thesis covered a broad range of task difficulty. The extremes were marked by Experiment 2 (fastest mean M=475 ms; central tendency μ = 414 ms) and Experiment 7 (slowest M=794 ms; μ=629 ms). The reaction times provide an idea of how different the experiments were with a relatively easy word-nonword discrimination in Experiment 2 compared to a relatively difficult task in Experiment 7. It could be argued that the lexical decision in Experiment 7 required a more thorough processing than in Experiment 2. The participants could have based their decision on less evidence in the

easy than in the more difficult experiment. One possibility of modelling this difference is by varying the thresholds in the decision channels of the model. A lower threshold would indicate a fast and to an extent sloppy response reflecting a situation like in Experiment 2. A higher threshold would reflect a more thorough processing mirroring Experiment 7.

The potential impact of the prime is greater when the decision threshold is low than when it is high. As mentioned earlier, the SCM averages the evidence across cycles and as Figure 4.6 illustrated the effect of the prime is limited to the early processing cycles. Thus, the effect of the early cycles is more pronounced in a fast than in a slow decision process. As a result, the impact of a prime decreases towards the end of a long decision process. That means that varying the decision thresholds is effectively manipulating the mixture of evidence from primes and targets. A similar prediction could be derived from the leaky accumulator model (Usher & McClelland, 2001) where evidence trickles out of the decision channels during the decision process. Thus, in a fast response the impact of the prime is greater than in a slow response and in the following simulations with the SCM the decision thresholds were fitted accordingly.

The weight of identifying a word target (y_{id}) was not altered in the following simulations compared to the value outlined in C. J. Davis (2010). This contrasts with the simulations in Chapter 4.2 which aimed at demonstrating the difference between homogeneous and selective inhibition whose impact is most apparent in summed lexical activity. Therefore y_{id} was reduced in Chapter 4.2. For the current simulations, it was assumed that the identification of a word provides strong evidence in favour of a *yes*-response. Thus, the y_{id} parameter as well as all other parameters apart from the inhibition between decision channels (λ, see 8.1) were set to the values outlined in C. J. Davis (2010). The weight assigned to summed lexical activity (y_{global}) and the decision thresholds (Θ_{yes}, Θ_{no}) were fitted to the empirical data.

8.2.3 Parameter search

A three-dimensional solution space was formed by the three parameters Θ_{yes}, Θ_{no} and y_{global}. A cost function describing the goodness of fit was formulated for estimating the best solution. The root mean square error of the model's predictions and the empirical data was formed. Since the SCM predicts the relative difference between two conditions (Davis, 2010), the congruency effect, the lexicality effect and the relative difference of the congruency effect in word and nonword targets was computed. Thus, the aim was minimising the function within the parameter space as shown in Formula 8.1.

$$\min_{0.0 \leq \Theta_{yes}, \Theta_{no} \leq 1.0 \wedge 0.0 \leq y_{global} \leq 2.0} \left(\frac{\sqrt{\Delta_{congruency\,effect}^{2} + \Delta_{lexicality\,effect}^{2} + \Delta_{relative\,congruency}^{2}}}{3} \right)$$

Formula 8.1: Cost function for fitting the SCM.

As outlined above the μ value of the RT distribution was used. In the following equations the μ refers to the empirical data and the M value to the prediction of the model. A list of all μ and M values is provided in Table 8.1 and the resulting values for the minuend and the subtrahend in Table 8.2 in the results section.

$$\Delta_{congruency\,effect} = \left(\frac{-\mu_{wc} + \mu_{wi} - \mu_{nc} + \mu_{ni}}{2} \right) - \left(\frac{-M_{wc} + M_{wi} - M_{nc} + M_{ni}}{2} \right)$$

Formula 8.2: Computation of the congruency effect for fitting the SCM.

$$\Delta_{lexicality\,effect} = \left(\frac{-\mu_{wc} - \mu_{wi} + \mu_{nc} + \mu_{ni}}{2} \right) - \left(\frac{-M_{wc} - M_{wi} + M_{nc} + M_{ni}}{2} \right)$$

Formula 8.3: Computation of the lexicality effect for fitting the SCM.

$$\Delta_{relative\,congruency} = \left(\frac{\mu_{wc} - \mu_{wi} - \mu_{nc} + \mu_{ni}}{2} \right) - \left(\frac{M_{wc} - M_{wi} - M_{nc} + M_{ni}}{2} \right)$$

Formula 8.4: Computation of the congruency effect relative to target lexicality.

Note, the conditions were abbreviated in all formulas: word congruent (wc), word incongruent (wi), nonword congruent (nc) and nonword incongruent (ni).

The cost function defines which of two points in the solution space represents a better fit. The value of the cost function could only be evaluated by simulating the experiment with the respective parameters and using the model's prediction in calculating the RMSE. Thus, looking at a fine grained grid of the three parameters and

the corresponding matrix of values in the cost function would have been computationally demanding. Although, the cost function makes the problem more manageable by adding the assumption that the goodness of fit is sufficiently described by the function, it does not alter the class of this NP complete problem. Thus, an heuristic algorithm was required for finding a good fit of the model. A number of algorithms have been developed for solving optimisation problems with several parameters. An important property of an appropriate algorithm is that the number of required values of the cost function is low, since computing these values is time consuming. The *Nelder-Mead simplex algorithm* (Nelder & Mead, 1965) was cited 9799 times ('Web of Knowledge - Citation Report', 2011) and is the most cited article in The Computer Journal ('Oxford Journals Reports — Most-Cited Articles as of August 1, 2011', 2011).

The Nelder-Mead simplex algorithm is based on a geometrical figure called *simplex* which is formed of $n+1$ points in an n-dimensional space. Each vertex is evaluated and the worst vertex is replaced with a better one. For achieving this, the worst point is reflected along the plane formed by the remaining vertices of the simplex. There are three possible results: a) the reflected point is the best solution of all vertices in the simplex, b) it is not the best solution, but better than the original point and c) it is even worse than the original point. In case of a) (best point) the simplex is expanded in the direction of the best fit. That is the reflection is performed not on a mirror, but on a mirror with a lens stretching the simplex towards greener pastures. For b) the algorithm finds the next worst point and starts again with a reflection. In case of c) the reflection is not performed and the simplex is contracted by moving the worst point towards the plane formed by the other vertices. This shrinks the simplex and reduces the search space. In the special case that the derived contracted point is even worse than its original, a contraction is rejected and the so-called multiple contraction is performed. In the multiple contraction step all vertices are moved towards the best vertex in the simplex. Once a step of the algorithm is completed, it starts again by identifying the new worst point in the simplex. The termination criterion can be defined as the minimum distance of the centre of the simplex between two steps. If

the transformation of the simplex does not result in a minimal difference, this indicates that the search space has been greatly reduced and the algorithm terminates.

The simplex algorithm is very effective in reducing the size of the search space, but it is prone to local minima of the cost function. Initial inspections of the solution space of Experiment 2 with two parameters supported the hypothesis that the cost function had several local minima. Those can be attributed to trade-offs between parameters, i.e. moving towards one extreme could make an unlikely option slightly more likely and form a local minimum. Specifically, a multiple contraction could move the figure from including the actual optimum to a shape where a local minimum traps the algorithm.

For coping with the local minima, the hybrid *Continuous Tabu Simplex Search* (CTSS) algorithm by Chelouah and Siarry (2005) was implemented. This algorithm combines the *tabu search* (Glover, 1989, 1990), which specifically handles local minima, with the effectiveness of the simplex algorithm (Nelder & Mead, 1965) in reducing the search space in a defined area. The tabu search explores the search space more widely. It starts at a random point in the solution space. For two or three dimensional problems that are considered here, a step in each possible direction is computed where the step size is usually between 0.1 and 0.2 in a normalised solution space. Thus, a geometric starlike shape is formed. The search moves towards the best point under consideration and adds all other points to a *tabu list*. The space around a point on the tabu list is not allowed to be the target of a step through space, because it is already known that one of its neighbours provides a better solution. The radius of the tabu list is usually smaller than the step size. If the central point of the star has the lowest cost of the points forming the star, it is a potential solution. In the original algorithm, the step size would be reduced until the optimum is found. In CTSS the central point and the most optimal outer vertices of the starlike figure are handed to the simplex algorithm for the more fine grained optimisation. In both accounts, this local solution is added to the *solution list*. Each solution point is surrounded by a tabu zone in the solution space which is larger than a single step size. The algorithm is restarted by choosing another random point in the solution space. By strictly avoiding the area of any previous solutions the search moves through the solution space until it either discovers another potential

solution or terminates in a cul-de-sac of prohibited areas. The termination criterion was set to five restarts without discovering another potential solution.

In searching for potential fits of the SCM the CTSS provided a set of potential solutions, but even though these solutions have been fitted by the simplex algorithm, an additional optimisation could be achieved by restarting simplex. This could be due to slightly too coarse termination criterion for the simplex algorithm with regards to the cost function. These changes were small and usually affected the second decimal. In sum, the CTSS provided a reasonable account of exploring the three dimensional solution space. The functionality of the SCM software was extended to include the simplex and the CTSS parameter search method.

8.3 Results of the simulations

The modified version of the SCM that incorporates selective inhibition and verification was used for the simulations. Three parameters Θ_{yes}, Θ_{no} and y_{global} were varied for fitting the model to the empirical data. As pointed out above the SCM predicts the relative differences between the conditions. Thus, RMSE values express the relation of the SCM's prediction to the relative effect sizes. For completeness, the absolute values are presented in Table 8.1 and the relative values with the respective RMSE in Table 8.2.

Table 8.1: Central tendency μ of the RT distribution of Experiment 2 through to 10 as a function of lexicality and the experimental condition and the respective predicted values derived by the SCM with selective inhibition and verification.

Lexicality	Word		Nonword	
Congruency	Congruent	Incongruent	Congruent	Incongruent
Experiment 2	414	430	431	457
Prediction	*48*	*65*	*65*	*91*
Experiment 3	431	442	460	482
Prediction	*36*	*46*	*65*	*86*
Experiment 4	441	467	454	473
Prediction	*43*	*63*	*54*	*73*
Experiment 5	445	455	502	510
Prediction	*72*	*82*	*128*	*137*
Experiment 6	451	460	496	500
Prediction	*69*	*80*	*110*	*119*
Experiment 7	501	500	629	624
Prediction	*159*	*161*	*286*	*285*
Experiment 8	484	496	583	586
Prediction	*74*	*83*	*172*	*172*
Prime typicality	High	Low	High	Low
Experiment 9	423	449	479	457
Prediction	*54*	*58*	*97*	*76*
Experiment 10	461	471	502	493
Prediction	*42*	*49*	*82*	*72*

Note that all values are rounded.

These results showed that the model was able to capture the effect of response congruency and typicality. In most cases the model provided a very good fit to the empirical data.

The results of the model simulations showed that the modified SCM was able to capture the lexicality and the congruency effect. The parameters that were found by the search algorithm are reported in Table 8.3. With the fitted parameters the predicted effect sizes were close to the empirically measured data. This was illustrated

Table 8.2: Comparison of the predictions by the SCM with verification and selective inhibition to empirical data of Experiment 2 to 10. The comparison was based on the relative effects of congruency and typicality on the central tendency of RT distribution. The RMSE is the measure of goodness of fit between the model's prediction and the empirical data. The table shows the value of the cost function at the best fit that was observed during the simulations.

Experiment	Congruency effect	Lexicality effect	Relative congruency effect	RMSE
Experiment 2	21.0	22.0	5.0	
Prediction	*21.5*	*21.5*	*4.5*	*0.4*
Experiment 3	16.5	34.5	5.5	
Prediction	*15.5*	*34.5*	*5.5*	*1.1*
Experiment 4	22.5	9.5	-3.5	
Prediction	*19.5*	*10.5*	*-0.5*	*2.5*
Experiment 5	9.0	56.0	-1.0	
Prediction	*9.5*	*55.5*	*-0.5*	*0.4*
Experiment 6	6.5	42.5	-2.5	
Prediction	*10.0*	*40.0*	*-1.0*	*2.7*
Experiment 7	-3.0	126.0	-2.0	
Prediction	*0.5*	*125.5*	*-1.5*	*2.2*
Experiment 8	7.5	94.5	-4.5	
Prediction	*4.5*	*93.5*	*-4.5*	*1.2*
	Typicality effect	Lexicality effect	Relative typicality effect	RMSE
Experiment 9	2.0	32.0	24.0	
Prediction	*-8.5*	*30.5*	*12.5*	*7.7*
Experiment 10	0.5	31.5	9.5	
Prediction	*-1.5*	*31.5*	*8.5*	*1.1*

Table 8.3: Comparison of the parameters produced in fitting the SCM with verification and selective inhibition to empirical data of Experiment 2 to 10.

Experiment	y_{global}	Θ_{yes}	Θ_{no}	Primes	Targets
Experiment 2	1.355	0.702	0.722	Easy	Easy
Experiment 3	1.134	0.448	0.664	Easy	Easy words, medium nonwords
Experiment 4	1.330	0.625	0.594	Easy	Medium words, easy nonwords
Experiment 5	0.382	0.335	0.887	Easy	Medium
Experiment 6	0.498	0.402	0.843	Medium	Medium
Experiment 7	0.309	0.893	0.943	Hard	Hard
Experiment 8	0.941	0.693	0.818	Easy	Hard
Experiment 9	1.905	0.753	0.728	Easy	Easy
Experiment 10	1.534	0.498	0.698	Easy	Medium

Compared to Experiment 2 the nonword targets were replaced by a more wordlike set in Experiment 3. In Experiment 4 the word targets were replaced, but the nonword targets were the same as in Experiment 2. Experiment 5 involved the nonwords from Experiment 3 and the words from Experiment 4. In Experiment 7 the nonwords were replaced with more wordlike items, i.e. "Hard" refers to the words from Experiment 4 and a new, very wordlike set of nonwords.

by the low RMSE values (see Table 8.2). There was one exception to this. In Experiment 9 the model underestimated the typicality-lexicality interaction (impostor effect). Table 8.1 shows that the SCM specifically underestimated this effect in word targets. As a result the interaction was underestimated and the magnitude of the main effect overestimated. The main effect refers to the average of the typicality effect in word and nonwords, which pointed into different directions. Thus, the similar effect sizes in word and nonword targets in empirical data resulted in a main effect of 2 ms (slightly larger effect in words), but the SCM's underestimation of the effect in word targets resulted in a prediction of -8 ms (larger effect in nonwords).

8.4 Discussion

The experiments in this thesis investigated the impact of task difficulty and prime typicality on the presence of response congruency and typicality effects. From the empirical data it was concluded that the task difficulty has an impact on the weight assigned to summed lexical activity (e.g., Grainger & Jacobs, 1996). Also, the amount of

evidence that was required for a lexical decision could have differed as a result of the experimental manipulation. Thus, the model was fitted to the data by varying the three parameters Θ_{yes}, Θ_{no} and y_{global}.

The results of the simulations showed that the SCM was able to capture the data in principal. Most of the fits to the data were very close, with the exception of Experiment 9. I discuss the impact of the typicality of nonwords, the task difficulty and the impact of the primes by comparing two pairs of experiments. Finally, I discuss the modifications to the SCM.

8.4.1 Typicality of nonwords

In Experiment 2 and Experiment 9, the word-nonword discrimination was similarly easy. But the y_{global} parameter was considerably lower in Experiment 2 (1.3) than in Experiment 9 (1.9). Nevertheless, both values were high compared to the original SCM parameters (0.4). In these two comparably easy experiments the decision thresholds were similar to each other. That means the amount of evidence that was collected before a response was triggered was similar for word and nonword targets. The SCM predicted a response congruency effect that was smaller in word than in nonword targets for Experiment 2. This prediction mirrored the empirical data with regards to the effect size in both word and nonword targets. But the model was unable to produce a typicality effect in word targets that was comparably large as in the empirical data of Experiment 9 which was due to an underestimation of the effect in word targets. Despite the higher value in y_{global}, the absolute response times of the model in congruent word trials were slower in Experiment 9 than in Experiment 2. This indicated that the advantage of the typical nonword primes was underestimated. Repeating the simulation of Experiment 2 with the parameters that were found for Experiment 9 revealed a congruency effect that was 1.5 cycles larger in word targets (and 46 cycles in nonword targets). By contrast the predicted typicality effect in Experiment 9 was 13 cycles smaller than in Experiment 2 despite a greater value of y_{global} and this highlighted that the model underestimated the lexical activity that was triggered by the high typicality nonwords. Thus, only this relatively high value of y_{global} enabled the model to predict a typicality effect at all. But a high value in y_{global} also contributes to faster

responses in incongruent word trials. The underlying reason is the relatively high setting of y_{global} also increases the impact of target related activity on the *yes*-channel. In this trade-off situation, this particular choice of y_{global} enabled the model to correctly predict the effect in nonwords and the interaction of lexicality and typicality. In summary, the effect size in empirical data was comparable between Experiment 2 and 9, but the model required a higher y_{global} in Experiment 9 in order to produce this effect. This showed that the SCM underestimated the lexical activity triggered by the high typicality nonword primes in Experiment 9.

The fact that the SCM underestimated the typicality of the high typicality nonword primes is also highlighted in the parameters of Experiment 10. The model provided a very good fit to the empirical data, but the value of y_{global} was the second highest of all simulations (1.5). That means this setting was still higher than in Experiment 2 even though the explicit task in Experiment 10 was more difficult than in Experiment 2. The simulations and previous computational work (Grainger & Jacobs, 1996) suggested that y_{global} should be lower in Experiment 9 than in Experiment 2. The fit to Experiment 10 provided here, showed that the SCM fitting algorithm compensated the underestimation of lexical activity in the primes by strongly emphasising the difference between high and low typicality primes through increasing y_{global}. Compared to Experiment 9 the typicality effect in the empirical data was smaller and the model was able to capture this effect. The small imbalance in the decision thresholds in the parameters for Experiment 10 was similar in all simulations that used this set of targets and could reflect that the word and nonword categories were not equally difficult for the model. In summary, the fitted y_{global} parameter in Experiment 10 supported the idea that the model underestimates the lexical activity from highly typical nonwords.

8.4.2 Task difficulty

In comparing Experiment 2, 3, 4, 5 and 7 the task difficulty steadily increased. The parameters of the model fits reflected this by decreasing the y_{global} constantly over the course of experiments. This was compatible with Grainger and Jacobs' (1996) findings. It is also interesting to note, that the fitted parameters showed an adjustment for the manipulation of word and nonword targets. From Experiment 2 to 3 the nonword

targets were made more difficult and parameters reflected this in a relative difference between the decision thresholds, i.e. Θ_{no} was larger than Θ_{yes}. Vice versa in Experiment 4, the word targets of Experiment 2 were replaced with more difficult items and Θ_{yes} was larger than Θ_{no}.

In Experiment 5, the step in y_{global} appeared very large compared to Experiment 3 and 4, especially since Experiment 5 essentially combined the targets of Experiment 3 and 4. This was also accompanied by a strong imbalance in the decision thresholds, more precisely Θ_{yes} was considerably lower than Θ_{no}. Even though the parameter choice might appear extreme compared to Experiment 3 and 4, the empirical difference was very pronounced as well. The lexicality effect in empirical data was about six times larger in Experiment 5 than in Experiment 4 and the congruency effect only half in size. Thus, reducing y_{global} enabled the model to accommodate these data. This result showed that the impact of using a combination of word and nonword targets that were more difficult to categorise than in Experiment 2 had a greater impact on the empirical data than just summing the steps from Experiment 2 to 3 and Experiment 2 to 4. But adding these steps together is exactly what the model does, unless the parameters are changed drastically as was shown in fitting the model.

The next step in task difficulty was from Experiment 5 to Experiment 7. Comparing the parameters of these two experiments shows that y_{global} was decreased even further and the decision thresholds were increased compared to Experiment 5. This choice reduced the response congruency effect and produced slower responses which mirrored the empirical data. But note, that there was no response congruency effect in Experiment 7 and the primes were also replaced compared to Experiment 5.

In summary, increasing the task difficulty was reflected in the parameters of the model by a reduced weight of summed lexical activity an larger decision thresholds. The comparison of Experiment 3, 4 and 5 indicated that participants experienced a strong non-linear increase of the task difficulty in Experiment 5. The fitted parameters revealed large differences between Experiment 5 and the previous experiments,

whereas Experiment 2, 3 and 4 could be accommodated by comparably similar sets of parameters.

8.4.3 Impact of replacing primes

The comparison of Experiment 5 and 6 and Experiment 7 and 8 can shed light on the impact of replacing the primes in the same set of targets. If the participants adapted to the explicit task only, then a difference between the parameter estimates of the model is not expected. The empirical data of Experiment 5 and 6 revealed only small differences of up to 10 ms between the same set of targets, where Experiment 6 that used the less informative primes consistently produced slower responses. Similarly, the responses in Experiment 7 using less informative primes were consistently slower than in Experiment 8. But in these experiments the difference was up to 46 ms in the same set of targets as a function of prime informativeness. These differences were most pronounced in nonword responses.

The differences in the empirical data implied that the parameters in the model fits differ for the same set of targets in order to accommodate the differences in the lexicality effect. In comparing Experiment 5 and 6, reducing the prime informativeness reduced the lexicality effect, whereas in Experiment 7 and 8 a reduced prime informativeness increased the lexicality effect. But in both cases the more informative primes resulted in a larger response congruency effect. The SCM could be adapted through the parameters to fit Experiment 7 and 8 more successfully than with Experiment 5 and 6. In Experiment 8 y_{global} was increased and by this means emphasising the lexical activity stemming from the primes. Also, the decision thresholds were reduced which enabled the model to produce faster responses with a smaller lexicality effect. Using the same approach would not result in a similarly good fit with Experiment 5 and 6, because the lexicality effect increased with more informative primes. Indeed, y_{global} was smaller with more informative primes in Experiment 5 compared to Experiment 6. Thus, the model overestimated the response congruency effect and slightly underestimated the lexicality effect in Experiment 6. Also the Θ_{yes} was smaller in Experiment 5 than Experiment 6 which contributed to capturing the larger lexicality effect in Experiment 5. The SCM can accommodate a

greater reliance on summed lexical activity if more informative primes are used with the same set of targets, but less so vice versa. The model readily captured a decreasing lexicality effect with more informative primes.

The finding that not only the congruency effect, but also the lexicality effect varied as a function of prime informativeness seems to support the hypothesis that participants can adjust unconsciously to what extent the primes are exploited (Bodner & Masson, 2001, 2003, 2004; Bodner et al., 2006). On the other hand, the primes were always equal in the proportion of useful (congruent/valid) and misleading (incongruent/invalid) stimuli. The primes were also formally and semantically unrelated to the targets. Thus, there was no obvious way of using the primes strategically nor was there an obvious exploitable relation between the primes and the targets.

An alternative account is related to the deep processing account (see Dehaene et al., 1998; Naccache & Dehaene, 2001). It could be argued that the participants applied the task instructions on the prime in all experiments. In informative primes some information was extracted from the prime that biased the lexical processing of the target, e.g. a word prime increased lexical activity in comparison to a nonword prime. The uninformative primes were processed in exactly the same way, but extracting information was either less successful or the result was less distinguishable between the word and nonword primes. So far, this would explain larger or smaller response congruency effects, but not the difference in the lexicality effects. These can be explained by the impact of the primes on the lexicon. The differences in response speed were more obvious in nonword than in word targets. This could be due to the fact that the nonword targets triggered relatively low levels of lexical activity, i.e. they are more prone to the influence of lexical activity induced by a prime than the word targets that produce activity themselves. During the decision process the *no*-channel has to work against activity in the *yes*-channel from a prime. With the informative primes the nonword primes were particular unwordlike and thus, a nonword response was not interfered by prime induced lexical activity. Thus, these nonword responses have a tendency for being faster than those related to an less informative (i.e. more wordlike) nonword prime. By contrast, the incongruent nonword trials receive slower

responses with informative primes because the word prime is more wordlike than in less informative primes. Depending on how typical the less informative primes were, the mean nonword response could be slightly faster or slower and thus, the lexicality effect could be larger or smaller in informative compared to less informative primes. This explanation crucially depends on how much lexical activity was triggered by a nonword prime. It was argued above that the SCM showed a tendency of underestimating the lexical activity that was triggered by a nonword and this could have contributed to the difficulty of the SCM of capturing this specific set of data.

In summary, it appears more likely that the participants extracted information about the typicality of the prime automatically (Dehaene et al., 2001; Naccache & Dehaene, 2001) and the difference in the lexicality effect can be to attributed to a different state of the lexicon rather than to a strategic effect. The SCM captured a smaller lexicality effect with increasingly informative primes more successfully than a pattern of increasingly informative primes paired with a larger lexicality effect.

8.4.4 Model modifications

For the purpose of simulating the experiments in this thesis the SCM had to be modified in two aspects. First, the lexical component of the model was adjusted. The homogeneous inhibition in the model was replaced by selective inhibition, as outlined in Chapter 4.2. Secondly, verification was introduced in the SCM. This mechanism resets the activity in the most active word node when the *yes*-decision threshold or its identification threshold was hit and there is a mismatch between the stimulus at letter level and the word node's associated template. Both mechanisms were active in the simulations in this chapter.

The selective inhibition in the lexicon was shown to be essential for predicting the typicality effect in word targets in Chapter 4.2. This modification in the lexical component did not impair the model's ability to predict inhibitory priming effects (see Chapter 4.3) which require inhibition between word nodes. The implementation of verification was necessary to enable the model to produce negative priming effects. The SCM was successful in predicting positive effects, but it also showed a small

negative effect on average in nonword targets in Experiment 7. Even though this negative effect was only small, it highlighted that the verification contributed to the simulations and that verification was an essential part in the simulations.

The SCM was also extended by a parameter search function which implements the simplex (Nelder & Mead, 1965) and the CTSS (Chelouah & Siarry, 2005) search algorithms. The parameter search above showed that the search algorithms worked and allowed fitting the model. It could be argued that the algorithm explored the parameter space faster than a manual search. Specifically because they did not require input or manual adjustments in the settings for the next simulation step allowing the parameter search to be run unsupervised. The simplex search is handy if a region of interest has already been identified, whereas the CTSS explores the parameter space more broadly. The results supported the idea that CTSS was effective in finding several minima in each simulated experiment. First, this showed that the parameter space was broadly explored and, secondly, that simplex alone could have been trapped by one of the local minima. With the combined search the SCM was enabled to find a good fit automatically.

8.5 Summary

The SCM was successful in predicting the typicality effects that were observed in Experiment 2 through to 10. The RMSE scores were very low indicating a very good fit, i.e. the predicted typicality and lexicality effect in cycles matched the empirical effect in ms very closely. The exception to this was Experiment 9 where the model underestimated the magnitude of the typicality effect. Furthermore, the results showed that modifications to the SCM, i.e. the introduction of selective inhibition and verification, increased the scope of the model. The parameter search algorithms that were incorporated into the software were also shown to be effective. Finally, the results of the simulations indicated that the observed typicality effects were the result of lexical activity, rather than motor activity or some other process that bypassed the lexicon.

The simulations in this chapter showed that a computational model of visual word recognition such as the SCM can explain the results obtained in Experiment 2 through to 10. But the SCM underestimated the lexical activity that was triggered by the highly typical nonword primes in Experiment 9 and 10. Also, the empirical data indicated that the impact of task difficulty is not a linear function of combining specific targets, thus the parameters in Experiment 5 (a combination of more difficult targets) differed strongly from those in Experiment 3 (more difficult nonword targets) and 4 (more difficult word targets compared to Experiment 2) in order to enable the model to stretch for accommodating the empirical effects. In general, all simulations showed that response congruency and typicality effects were the result of summed lexical activity and it was shown in Chapter 4 that selective inhibition was necessary to produce an increase in lexical activity.

9. General Issues

9.1 Summary empirical findings

9.1.1 Response congruency effect

A response congruency effect in masked primed lexical decision refers to a situation where word responses are facilitated by formally unrelated word primes compared to unrelated nonword primes, and nonword responses are facilitated by nonword primes. Norris and Kinoshita (2008) argued that response congruency effects would not emerge in lexical decision which is compatible with their own and other data reported in the literature (Perea et al., 1998, 2010). On the other hand, the possibility of obtaining a response congruency effect appears to be a critical prediction of an activation framework such as the Spatial Coding Model (SCM, Davis, 2010). Klinger et al. (2000) reported evidence for a response congruency effect in lexical decision. However, in this experiment the stimuli were repeated and thus, the congruency effects could be attributed to stimulus-response mappings that participants learnt during the experimental session (Damian, 2001), rather than to process occurring in the lexicon. The experiments reported in this thesis attempted to investigate the existence of response congruency effects in the lexical decision task more systematically.

The results of Experiment 2 clearly showed that response congruency effects can emerge in masked primed lexical decision. The stimuli in this experiment were not presented repeatedly and thus, participants were not able to form stimulus-response mappings. It was concluded that the observed response congruency effect in masked primed lexical decision was the result of lexical activity induced by word primes that facilitated a *yes*-response in word targets and interfered with a *no*-decision in nonword targets. Experiment 2 comprised a relatively easy word-nonword discrimination task and it could potentially be argued that this was not comparable to a standard lexical decision task. The difficulty of the explicit word-nonword discrimination task was systemically increased in Experiments 3 to 5. All three experiments revealed robust

response congruency effects and provided evidence that these effects can emerge in masked primed lexical decision.

The experiments using an increasingly difficult explicit task also tested whether response congruency effects are more prevalent in fast responses (see 1.4.2.3). The experiments in this thesis showed that the response congruency effect is robust with an increasing task difficulty and in slower responses. Figure 9.1 shows the response congruency and typicality effects of all experiments as a function of the mean RT. The graph shows that the effect tends to be smaller in an experiment with slower responses than in an experiment with comparable fast responses. This is compatible with the decreasing effect sizes that were reported with an increasingly difficult explicit task (e.g., in Experiment 2 to 5). However, the smaller response congruency effects were also partly due to a mixture of positive and negative effects. This was shown using a more fine-grained analysis in deciles. That means that response congruency effects were observed with slow responses. These effects were small in word targets and tended to be negative in nonword targets (see Figure 9.1). In summary, the response congruency effects reported in this thesis were influenced by the task difficulty and average response speed, but they did not depend on these factors.

An important difference between Experiment 2 to 5 and data reported in the literature were the primes. In Experiment 2 to 5 the primes were clear exemplars of

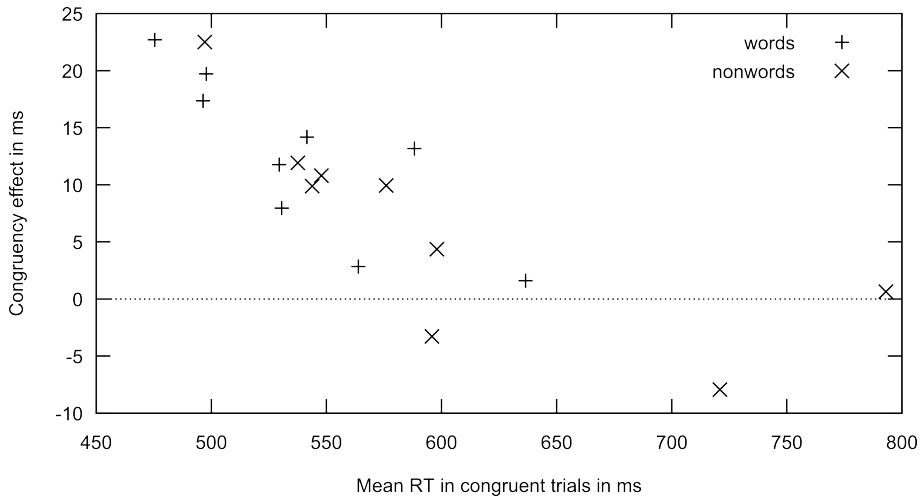

Figure 9.1: Plot of the response congruency effect in Experiment 2 through to 8 and the typicality effect in Experiments 9 and 10. The mean RT reflects the congruent trials in Experiments 2 to 8. In Experiments 9 and 10, the high typicality priming condition in word targets and the low typicality priming condition in nonword targets was used as a basis.

their categories. That means the word primes were high in frequency and orthographically typical (e.g., *royal*, *thing*) whereas the nonwords had no orthographic neighbours (N; Coltheart et al., 1977) and were orthographically atypical (e.g., *dvnel*, *oyizi*). It was hypothesised that the primes in these experiments were very informative with regards to the lexical decision task and that they biased the decision process towards a *yes*- or *no*-response respectively. This was in contrast to experiments in the literature (Norris & Kinoshita, 2008; Perea et al., 1998, 2010) where the word and nonword primes were very similar to each other with regards to their typicality and *N*. Additionally, Norris and Kinoshita (2008) used nonword primes that were also formally very similar to specific word primes such as *chill* and *chilk*, *floor* and *floop*. Concluding from this, it was assumed that the informativeness of the primes, i.e. the relative difference in typicality between the word and nonword primes, was the distinctive property of the experiments in this thesis and the literature.

This hypothesis was tested in the following experiments. In Experiment 6, the primes were less informative compared to Experiment 2 to 5. That means compared to the previous experiments the word primes were lower in frequency and less typical (e.g., *rifle*, *solid*) and the nonword primes were somewhat more typical (e.g., *dulew*, *saulp*). This reduced the relative difference between the prime categories and thus, the informativeness of the primes with regards to the word-nonword discrimination task. The results of Experiment 6 showed that the response congruency effect was reduced in word targets and diminished in nonword targets. This suggested that a stronger manipulation of prime informativeness could result in the absence of response congruency effects and resolve the empirical discrepancy with the literature.

In Experiment 7 the prime informativeness was further reduced compared to Experiment 6 and the targets for the explicit word-nonword discrimination were made more difficult to categorise. The word items (e.g., *dwarf*, *zebra*) were orthographically less typical than the nonword items (e.g., *grome*, *sheme*). The results of Experiment 7 showed that there was no response congruency effect in either word or nonword targets. Since both primes and targets were replaced compared to Experiment 6, another experiment was required to unambiguously demonstrate that the prime

informativeness was the distinctive factor underlying the differences between Experiment 2 to 6 and the previous literature. Experiment 8 used the same targets as Experiment 7, but the more informative primes from the previous experiments. The results of this experiment revealed a response congruency effect. This showed that within the same targets a response congruency effect can be present or absent as a function of the prime stimuli. The use of more informative primes explained the empirical discrepancy between the response congruency effects reported in this thesis and the absence of such effects in the literature.

9.1.2 Typicality effect

The results of Experiment 2 to 8 showed that the response congruency effect critically depends on prime informativeness, specifically the relative differences in typicality between the word and the nonword primes. This implied that the presence or absence of a congruency effect was not determined by the lexicality of the prime, but by the informativeness of the prime with respect to a word- or a nonword-response. The orthographic typicality which was the main factor in manipulating the prime informativeness is independent of the prime's lexicality. The required manipulation in the stimuli can be achieved more easily in nonwords, because words always resemble actual language and are necessarily wordlike and typical to some extent. Thus, sets of typical nonwords and atypical nonwords were generated by using the software described in Chapter 2.6. In Experiment 9, a similar set of targets as in Experiment 2 was used for creating an easy lexical decision task, but the primes were a set of typical nonwords and a set of atypical nonwords. The results revealed that typical nonword primes facilitated *yes*-responses compared to atypical nonwords and vice versa atypical nonword primes facilitated *no*-responses compared to typical nonword primes. These results supported the argument that the response congruency effect in the earlier experiments was in fact driven by the typicality of the primes. Since the targets in Experiment 9 formed a relatively easy word-nonword discrimination task, the experiment could be criticised on the grounds of having used items that made the explicit lexical decision task too easy for participants. Thus, the experiment was repeated with a more difficult explicit task. The results of Experiment 10 were compatible with Experiment 9 in showing that typical nonword primes facilitated *yes*-

responses compared to atypical nonword primes. There was not a significant effect in *no*-responses, though there was a strong tendency in the same direction as in Experiment 9. In summary, Experiment 9 and 10 revealed an interaction of prime typicality and target lexicality, where typical primes facilitated *yes*-responses compared to atypical primes and vice versa atypical primes facilitated *no*-responses compared to typical primes. These findings supported the idea that the response congruency effects reported above were in fact typicality effects.

On the basis of this experimental evidence it can be concluded that the reasons the experiments previously reported in the literature did not obtain response congruency effects (Norris & Kinoshita, 2008; Perea et al., 1998, 2010) is that they did not use sufficiently informative primes. Nevertheless, the observed effects are better characterised as typicality effects rather than congruency effects, given that the typicality manipulation was effective in the experiments independently of the lexicality of primes. The findings are compatible with Woollams et al.'s (2011) findings that orthographic typicality is independently processed from lexicality. The effect of prime typicality in Experiment 10, that required participants to identify the targets, also suggested that typicality contributed to the word recognition process.

9.1.3 Negative typicality effects

The experiments reported here showed a robust typicality-lexicality interaction, whereby there was an advantage in RT and/or accuracy when a word target was preceded by a highly typical prime compared to an atypical prime and vice versa in nonwords. Nevertheless, in some experiments a typicality effect that went in the opposite direction was observed in the latter part of the RT distribution, i.e. highly typical primes facilitated *no*-responses and atypical primes facilitated *yes*-responses. The standard analysis of the empirical data was too coarse for detecting these negative priming effects. Only the application of more fine-grained methods such as forming vincentiles (Vincent, 1912) and fitting the data to an Ex-Gaussian distribution (Balota et al., 2008; S. Brown et al., 2008; S. Brown & Heathcote, 2003; Heathcote et al., 2004) allowed the detection of these negative effects in the data. One consequence of failing to detect the presence of these negative effects would have been an underestimation

of the positive effect that was present in the rest of the RT distribution. Another consequence would have been the failure to detect evidence supporting a reset of activity in word nodes.

In the data, positive and negative typicality/congruency effects occurred together in the same experiment. Specifically, the positive effect was more present in fast responses and the negative effects occurred in the slowest part of the RT distribution. For example, in nonword responses in Experiment 3 and in word responses in Experiment 10 a positive effect was found in fast responses and in the final deciles a negative effect was observed. Also, the congruency effect and the typicality-lexicality interaction faded out towards the end of the RT distribution, e.g. in nonword responses in Experiment 9 and word responses in Experiment 5. But this could also indicate that there was an increasing number of responses showing a negative effect towards the end of the distribution, rather than no effect at all. The fading out is different from the negative effects in that there was no clear boundary between positive and negative effects as in Experiment 3, but rather a mixture of both. These findings were explained by assuming a verification mechanism that can lead to the reset of activity in word nodes.

A verification mechanism that results in a reset of a word node is different from leakage (see Usher & McClelland, 2001). If the variation in the size of the effect over the course of the RT distribution were simply the result of leakage in the decision channels, the effect would slowly diminish towards the end of the RT distribution. But the empirical data implied a more abrupt mechanism that reversed the direction of the typicality and congruency effect. Thus, leakage could not explain the observed pattern. It is also important to note, that an effect that is negative still indicates an effect of typicality priming. In a different priming paradigm, Boy and Sumner (2010) showed that negative effects can only emerge if there is at least a potential for a positive effect, i.e. a negative priming effect is essentially due to the same processes as a positive effect and some additional process (inhibition, reset or verification) that is initiated with a delay relative to other processes. Boy and Sumner (2010) suggested that negative effects emerge if a response option hits an inhibition threshold, but not the response

threshold, i.e. the inhibition process is only started once the inhibition threshold was hit, but not when the stimulus was presented. In the SCM (Davis, 2010) a verification process was added that is initiated after the identification or *yes*-response threshold has been hit and thus, its onset is delayed relative to other processes. In summary, negative priming effects are the result of the same processes as positive priming effects, but with an additional mechanism that inhibits activity on the basis of the failure of a verification process. In the data presented in this thesis, negative effects were typically accompanied by positive effects which is compatible with this suggestion. Also, fine-grained analysis methods were required to detect negative effects, specifically if a mixture of positive and negative effects could have covered an actual effect.

9.1.4 Relation of typicality priming to other priming effects

It is interesting to compare the typicality priming effects observed here with unrelated primes to priming effects that have previously been studied for related primes. One noteworthy difference is the impact of target lexicality. In Experiment 2 through to 6 and 8 to 10, the effect of a congruency or typicality manipulation in the primes was reported and it was shown that the effects are absent if the primes are not sufficiently informative, i.e. if the typicality manipulation was not sufficiently strong. In Experiment 2 and 9 where the most pronounced effects were observed, the absolute effect was greater than 20 ms in word and nonword targets. The relative effect size in Cohen's *f* was greater than 1.00 indicating a large effect size. Specifically, the effect in nonword targets was unusually large. In a review Forster et al. (2003) estimated the effect from identity priming in nonword targets in lexical decision at about 8 ms. This contrasts with the identity priming effect in word targets which is one of the largest and most stable priming effects with an effect size that is as large as the prime presentation duration and occasionally even greater than that (Forster et al., 2003). That is, there is a very strong interaction of target lexicality and the identity priming effect. By contrast, the effect sizes of the congruency effect in Experiment 2 and the typicality-lexicality interaction in Experiment 9 were comparable in word and nonword targets. Specifically, the effect in nonword targets was about three times as large as in identity priming even though there was no letter overlap between prime and target.

The comparable magnitude of the congruency/typicality effects for word and nonword targets implied that the mechanism underlying these priming effects are different from those underlying identity priming and form priming. In particular, identity and form priming effects appeared to be the result of processing related to a specific word node, whereas the congruency and typicality effects reported here were attributed to general lexical activity independent of the target representation. This results in congruency/typicality effects in both word and nonword targets.

Another respect in which the congruency effect differs from form priming effects is the effect of word frequency. The congruency effects reported here appeared to be independent of the lexical frequency of the target and the relative frequency of prime and target. In contrast, form priming effects depend on the relative frequency of prime and target. For example, Grainger (1990) showed that using a neighbour prime that is higher in frequency than the target resulted in an inhibitory priming effect in lexical decision, but if the prime was lower in frequency there was no effect (though there was a facilitatory tendency). In a similar experiment, C. J. Davis and Lupker (2006) showed that a form related nonword prime facilitated a lexical decision in word targets whereas a related word prime showed an inhibitory effect relative to unrelated primes of the same lexicality. The size of this effect was attenuated in high frequency word targets compared to low frequency words in C. J. Davis & Lupker's (2006) data. In contrast, the current study of typicality revealed the largest effects with high frequency targets. The neighbourhood effects were attributed to inhibition between the prime and the target word nodes. But the congruency effect and typicality-lexicality interaction were attributed to word nodes that are not related and crucially, do not have inhibitory connections. Thus, it does not matter whether the target or the prime is of relatively higher frequency, because the effect of inhibition between prime and target is essentially absent. Furthermore, the absence of inhibitory connections between prime and target was crucial for simulating the data (see Chapter 4).

Both the comparison to masked identity priming and neighbour primes showed that the congruency effect and the typicality-lexicality interaction were not the result of local processes in the lexicon but of global processes. As a result the target lexicality,

the target frequency and the relative prime-target frequency were less critical. These empirical differences underline that the congruency effect and the typicality-lexicality interaction were the result of a different mechanism compared to priming effects in formally related primes. The form priming effects relied on the identification of a specific item whereas the effect of congruency and typicality was attributed to summed lexical activity. Differentiating between these processes is also compatible with the fMRI data by Woollams et al. (2011) who found that lexical identification and the typicality of letter strings activated different brain areas.

9.1.5 Theoretical accounts of response congruency

The Experiments 2 through to 8 used response congruency effects in lexical decision for investigating visual word recognition. Thus, it was important that these effects could not be attributed to other processes that were not related to word recognition. Damian (2001) presented evidence that response congruency effects can emerge if the primes have been used repeatedly as targets before. This finding was critical for Klinger et al.'s (2000) work where a response congruency effect was observed, but it could be attributed to stimulus-response mappings that the participants learnt during the experimental session. In order to avoid this criticism, the experiments in this thesis presented the targets only once. Furthermore, the congruency effect was observed in a number of experiments that used primes and targets from disjoint sets of items. In Experiment 9 and 10 a typicality-lexicality interaction was observed which cannot be attributed to stimulus-response mappings, because all primes were nonwords and thus, associated with the same response.

The deep processing account (Dehaene et al., 1998; Naccache & Dehaene, 2001) assumes that the task instruction are applied to the prime and the target and both stimuli are truly evaluated with regards to their meaning in the task. It was also argued that the prime and the target are processed by the same mechanisms, though the prime related processing is incomplete. This contrasts with the stimulus-response mapping (Damian, 2001) and the action trigger account (Kunde et al., 2003), where it was assumed that a response can be triggered without processing the meaning of the stimulus. This general principle is compatible with the SCM (Davis, 2010) as all primes

undergo the same processes as the target and their processing is incomplete. In the deep processing account the locus of the congruency effect is at the response level. An incongruent prime triggers a small response readiness potential and the target response interferes with the prime related potential. In a congruent trial the prime and the target related response coincide and thus, the target response can benefit from the prime. This is also compatible with SCM. Although the SCM performs a reset of the decision channels at target onset, the early advantage of congruent trials is preserved in the decision channels. Thus, it could be argued that both accounts are similar in that the locus of the congruency effect is at the decision/response channels. A critical assumption in the deep processing account is that the response channel with the prime related response is activated, specifically in Experiment 9 and 10 this would not predict the observed typicality effects. But it could be argued that the word recognition system is tricked into falsely adding activity to *yes*-channel even though the stimulus is a (very typical) nonword. Thus, the deep processing account can accommodate the findings presented in this thesis. In summary, the deep processing account provides a framework that can account for the response congruency and typicality effects that were reported in this thesis. But it does not provide insight into the processes that are involved in activating the *yes*- or the *no*-channel as it is not a domain specific model.

The semantic feature overlap account (Quinn & Kinoshita, 2008) predicts response congruency effects on the basis of features that can be activated by the prime and the target stimulus. If there is an overlap between the prime and the target related set of features a facilitatory effect is predicted. The account assumes that in a narrow category where only a few features are important to monitor, a priming effect is predicted in both exemplar and nonexemplar responses. That means if words are associated with particular features that can describe the letter string, such as *pronounceable*, *legal* and orthographically *typical*, a narrow category of meta-features could be formed. Quinn and Kinoshita (2008) showed that nonexemplar words (e.g., *mind*) that share semantic features with a category (e.g., *body parts*), trigger similar effects as an actual category member (e.g., *head*). These words were referred as impostors and in Experiment 9 and 10 a similar concept was applied in creating nonword primes that were very wordlike and tricked the recognition system in

assuming this was a word. In summary, the overlap hypothesis provides a framework of response congruency that relies on features of the stimuli, but not a specific account of how these features are activated or formed. Nevertheless, the impostor concept provided an interesting input into the experimental work.

In summary, the theoretical accounts provided a framework of potential sources of response congruency effect that should be avoided (e.g., stimulus-response mappings). Also, these accounts provided a framework for potential effects, such as the impostor effect. Such an account is domain unspecific and after being transferred into the domain visual word recognition provided insight in the early processing stages.

9.2 Modelling response congruency and typicality effects

The simulations with the Bayesian Reader (Norris, 2006; Norris & Kinoshita, 2008) showed that this model cannot account for the response congruency effect. In the Bayesian Reader all traces of an unrelated prime are wiped off the lexicon as soon as the target was presented. Another type of model that was not tested in the simulations is the search model. For example, the entry opening model (Forster, 1985, 1999; Forster et al., 1987, 2003) has no means of capturing the effect of an unrelated prime on the lexical decision to the target. If a word prime initiated opening its corresponding entry in the lexicon, there is no way in which this could benefit opening another unrelated entry corresponding to a word target. Only an activation model, such as the Spatial Coding Model (SCM; Davis, 2010) seems to be able to account for the effects.

In modelling the response congruency effect and typicality-lexicality interaction using the SCM (Davis, 2010) it was shown that the impact of summed lexical activity can explain the findings. Importantly, the modelling showed that the lexical component could only account for the findings if selective inhibition was used. When the model used homogeneous inhibition as in the Interactive Activation Model (McClelland & Rumelhart, 1981), the SCM was unable to account for the observed response congruency effects in word targets. The reason was that activity induced by the prime word interfered with activity induced by the target word and this resulted in a null effect of response congruency in word targets. By using selective inhibition this

interference could be avoided and the model showed an increased level of lexical activity in congruent word trials compared to incongruent word trials. It was argued that selective inhibition is more plausible, because in this account only those word nodes interfere with each other that are related. This limits the competition for the best match for the stimulus to a set of reasonable candidates. In contrast, homogeneous inhibition implies that activity in each word node in the lexicon is actively suppressed whenever a letter string is processed. Selective inhibition is more restrictive in its application but it still allowed the model to capture data that requires inhibitory connections, this includes the inhibitory priming effect of word neighbours in word targets (Davis & Lupker, 2006). In summary, the implementation of selective inhibition increased the scope of the SCM. Since the lexical component in the SCM is very similar to the one in the Interactive Activation Model (McClelland & Rumelhart, 1981), the use of selective inhibition could generalise to other models that are based on this model, e.g. MROM (Grainger & Jacobs, 1996).

It is interesting to note, that the SCM predicted a typicality priming effect in nonwords independently of whether inhibition was homogeneous or selective. By comparison with word targets, the nonword targets triggered relatively little lexical activity. With homogeneous inhibition the activity stemming from the word targets resulted in inhibition of the prime word node and eventually in the false prediction of a null effect. Thus, the fact that the nonwords only triggered little lexical activity made the congruency effect in nonword targets immune to the choice of homogeneous or selective inhibition.

For capturing negative priming effects a reset mechanism was implemented in the SCM. This mechanism is similar to the verification mechanism in the activation-verification model (Paap et al., 1982). When this verification fails, a reset of the relevant word node helps to avoid a false *yes*-response to nonword targets that have high frequency word neighbours. In the present implementation, the verification is triggered if the *yes*-channel hits the decision threshold or a word node hits the identification threshold. If the word node is not a match for the stimulus, the activity of that node is reset. This reduces the input to the *yes*-channel and eventually enables the

model to correctly produce a *no*-response. That means the verification process helps in responding correctly. Since the verification can be triggered by hitting the *yes*-decision threshold, the prime can influence when this verification is triggered. A typical prime contributes to the activity in the *yes*-channel so that in turn the threshold is reached faster than with an atypical prime. Thus, the verification is triggered faster as well and in a nonword target this results in a negative typicality effect. For example, a very typical prime (e.g., *ifter*) would increase activity in the *yes*-channel to a greater extent than an atypical prime (e.g., *miqyd*) and thus, the threshold would be hit faster with the typical prime. A typical nonword target triggers further activity in the lexicon and the *yes*-channel. Consequently, for nonword targets, verification is most likely to occur (and occurs earlier) when a typical target is preceded by a typical prime. This in turn increases the likelihood of reset in typical relative to atypical primes, and reset leads eventually to a faster correct *no*-response, explaining the negative typicality effect. The role of the verification process in explaining the negative priming effects in the empirical data adds further plausibility to the idea of such a mechanism.

In summary, simulating the empirical data showed that selective inhibition is more appropriate than homogeneous inhibition and added evidence in favour of a verification process. The simulations also suggested that only an activation model could capture the congruency effect and the typicality-lexicality interaction.

9.3 Measuring orthographic typicality

The results of all experiments presented in this thesis showed the effect of a global measure of typicality on the word recognition process. Woollams et al. (2011) found evidence in fMRI data that typicality is represented differently from lexicality, but also that typicality is processed earlier than the identity of words. Various methods of measuring typicality have been suggested in the literature, e.g. OLD20 (Yarkoni et al., 2008) and transition frequencies of subsyllabic elements (Keuleers & Brysbaert, 2010). In Chapter 2, I introduced a measure of orthographic typicality based on trigram transition probabilities. Despite not using any explicit phonological information like Wuggy (Keuleers & Brysbaert, 2010) the algorithm produced almost all pronounceable nonwords (99%, see Appendix C) compared to 80% in WordGen (Duyck et al., 2004).

Also, the algorithm has a generic formulation that could be applied to a phonetic code or any other string of characters. The typicality of a string is expressed in standard deviations and is independent of the specific language. Also, the distribution of values that stems from analysing the CELEX lexicon resembles a typical RT distribution with fewer typical items (fast responses) and a longer tail towards atypical items (slower response). The evidence showed that this measure is an effective estimate in nonword targets. By specifying a specific layer of typicality, the generated nonword items vary in their wordlikeness and in turn in their ease of pronounceability (J. Humphreys, 2008). Thus, nonwords for reading ability tests could be generated with different levels of difficulty. This could be important if patients require a regular testing of their progress, but the same items cannot be used again because patients could just remember them. Existing test batteries, such as PALPA (Kay, Lesser, & Coltheart, 1996) and LeMo (De Bleser, Cholewa, Stadie, & Tabatabaie, 1997), could be extended. The generic implementation of *OT3* also allows the use in other languages.

In Experiment 1 low typicality items resulted in very similar RT independently of their pronounceability (but see Rubenstein et al., 1971, 1975). This indicated that pronounceability did not show an effect per se, although the effect of typicality was larger within pronounceable than within unpronounceable nonwords, i.e. the pronounceability could have contributed to the wordlikeness of a nonword above and beyond its typicality. Further evidence that typicality influences the word recognition process was obtained in the masked priming experiments especially Experiment 9 and 10.

9.4 What is the purpose of selective inhibition?

All experiments in this thesis were lexical decision experiments in the lab and most of them used the masked priming paradigm. Thus, it could be argued these experiments are not resembling natural reading conditions. Nonetheless, the results showed a basic principle of word recognition. It was demonstrated in the modelling sections that a selective inhibition account is more appropriate for simulating the data. It could also be argued that selective inhibition is more economic than homogeneous inhibition. One could assume that every transmitted signal in the lexicon is associated

with cost (chemical energy). Using homogeneous inhibition, presenting a word to the recognition system would trigger inhibitory signals being sent to about 50,000 word nodes. The number of connections can be reduced by implementing homogeneous inhibition as a function of summed lexical activity (see McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982) compared to laterally connecting each word node with every other word node in the lexicon. But crucially, the total number of transmitted inhibitory signals is not affected by changing the emitting source. If the amount of transmitted signal is at least the amount of received signal, then homogeneous inhibition requires about 50,000 inhibitory signals being sent from a word node or a sum function to each word node. By contrast, this number in considerably smaller in a selective inhibition account. In this account it is sufficient that each word nodes has an inhibitory connection to word nodes that are similar and thus, take part in the competition for the best match. In the simulations in Chapter 8 it was assumed that his number is limited to the 30 most similar word nodes. Thus, the amount of inhibition in a selective account is 1,000 times smaller than in a homogeneous account.

The experiments provided very clear indication that unrelated word nodes are not affected by inhibition, i.e. are not actively suppressed by the presentation of unrelated words or nonwords. In contrast to lexical decision experiments, words do not appear in isolation in normal reading. Most written documents enable the reader to potentially process more than one word. In a review, Rayner (1998, 2009) argued that words in the parafovea are processed with regards to their form (e.g., letters, morphology, phonology) but not semantically. That means there is evidence that more than one word could be processed at a time in the form lexicon. Ideally, these words do not interfere with each other, but are helpful for the reading process. With homogeneous inhibition the activity in two word nodes would cause strong interference and finally the suppression of activity in one of them. Thus, the prediction of homogeneous inhibition is that activity in the whole lexicon is actively suppressed as soon as a letter string is processed. That means as soon as one starts reading the lexicon is shifted into a dim state. By contrast, selective inhibition allows two active word nodes and inhibition is limited to word nodes that could cause confusion with the actual stimulus.

Another way of enabling the recognition system to process words in fast succession is a reset mechanism. Grainger and Jacobs (1999) argued that the activation in a word node could be eliminated if a mismatch occurs. This resembles the reset in adaptive resonance theory where a mismatch with the input resets the respective representation (Carpenter & Grossberg, 1987). The idea of a reset mechanism was implemented in the SCM as well in order to account for negative priming effects. In this account the activity that is typically triggered by the prime was not sufficient to make it a potential goal for a reset mechanism. Thus, the reset is not hindering the response congruency effects. But with homogeneous inhibition these effects are hard to obtain as the simulations in Chapter 4 showed. This implies that it was still necessary to use selective inhibition.

The experimental work presented in this thesis provided clear experimental evidence in favour of a selective inhibition account and constraints computational models of word recognition.

9.5 What is the purpose of orthographic typicality?

The experiments in this thesis were all concerned with the orthographic typicality of letter strings and it was argued that typicality is reflected in summed lexical activity. The results indicated that the typicality has an impact on the processing of words in isolation which is compatible with other results in the literature (Hauk et al., 2006; Woollams et al., 2011). Thus, it is important to know in which way this information is of use in the normal reading process.

Woollams et al. (2011) argued that typicality is processed earlier than the identity of a word. Thus, in reading the typicality of letter string could guide the planning of saccades. Eye tracking experiments have shown that lexical identification can continue while a saccade to the next word is being performed (Irwin, 1998; Yatabe, Pickering, & McDonald, 2009). Furthermore, the effect of word frequency also spilled over to the next word and provided evidence that word identification continues even after another word is being fixated (Kennison & Clifton, 1995; Yatabe et al., 2009). This implies that the word was not identified, but rather the word recognition system had some

indication that the stimulus can be identified soon. The SWIFT model of saccadic eye movements (Engbert, Nuthmann, Richter, & Kliegl, 2005; Richter, Engbert, & Kliegl, 2006) assumes that word difficulty feeds into the planning process which was expressed as a function of the word's frequency and its predictability. Both measures are dependent on the identity of a word and thus, are only available once the word is known but not in an on-the-fly process. But the frequency of a word influences how quickly its activation in the lexicon raises and it could be argued that the growth rate of summed lexical activity could provide a sufficient approximation for the planning process. In this thesis, it was argued that the orthographic typicality of a letter string is strongly linked to the summed lexical activity. The effect of impostor primes (very wordlike nonwords) in Experiment 9 and 10 was comparable to word primes in the previous experiments and it could be hypothesised that impostor nonwords cause the word recognition system to erroneously allow a saccade even though they might require some more processing to be identified as a very typical unknown stimulus.

The literature on eye tracking suggests that words in the parafovea are processed to some extent but not identified (Rayner, 1998, 2009). This could suggest that quickly extractable information such as typicality is used in planning landing points or in skipping highly typical function words.

The experiments in this thesis highlighted the importance of typicality in the word recognition process. The simulations showed that selective inhibition rather than homogeneous inhibition is required in order to capture these findings in computational models.

References

Abrams, R. L. (2005). Unconscious processing of multiple nonadjacent letters in visually masked words. *Consciousness and Cognition*, *14*, 585-601. doi:10.1016/j.concog.2005.01.004

Adelman, J. S., & Brown, G. D. A. (2008a). Modeling lexical decision: The form of frequency and diversity effects. *Psychological Review*, *115*(1), 214-227. doi:10.1037/0033-295X.115.1.214

Adelman, J. S., & Brown, G. D. A. (2008b). Postscript: Deviations from the predictions of serial search. *Psychological Review*, *115*(1), 228-229. doi:10.1037/0033-295X.115.1.228

Adelman, J. S., Brown, G. D. A., & Quesda, J. F. (2006). Contextual Diversity, Not Word Frequency, Determines Word-Naming and Lexical Decision Times. *Psychological Science*, *17*(9), 814-823. doi:10.1111/j.1467-9280.2006.01787.x

Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(5), 802-814. doi:10.1037/0278-7393.15.5.802

Andrews, S. (1992). Frequency and neighborhood effects on lexical access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(2), 234-254. doi:10.1037/0278-7393.18.2.234

Andrews, S. (1996). Lexical retrieval and selection processes: Effects of transposed-letter confusability. *Journal of Memory and Language*, *35*(6), 775-800. doi:10.1006/jmla.1996.0040

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Bulletin of the Psychonomic Society*, *4*(4), 439-461. doi:10.3758/BF03214334

Andrews, S., & Heathcote, A. (2001). Distinguishing common and task-specific processes in word identification: A matter of some moment? *Journal of*

Experimental Psychology: Learning, Memory, and Cognition, *27*(2), 514-544. doi:10.1037//0278-7393.27.2.514

Ansorge, U., & Neumann, O. (2005). Intentions Determine the Effect of Invisible Metacontrast-Masked Primes: Evidence for Top-Down Contingencies in a Peripheral Cuing Task. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(4), 762-777. doi:10.1037/0096-1523.31.4.762

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390-412. doi:10.1016/j.jml.2007.12.005

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1995). The CELEX Lexical Database. Release 2. Linguistic Data Consortium, University of Pennsylvania, Philadelphia.

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(3), 340-357. doi:10.1037/0096-1523.10.3.340

Balota, D. A., & Chumbley, J. I. (1990). Where are the effects of frequency in visual word recognition tasks? Right where we said they werep Comment on Monsell, Doyle, and Haggard (1989). *Journal of Experimental Psychology: General*, *119*(2), 231-237. doi:10.1037/0096-3445.119.2.231

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, *133*(2), 283-316. doi:10.1037/0096-3445.133.2.283

Balota, D. A., Tse, C.-S., Hutchison, K. A., Spieler, D. H., Duchek, J. M., & Morris, J. C. (2010). Predicting conversion to dementia of the Alzheimer's type in a healthy control sample: The power of errors in stroop color naming. *Psychology and Aging*, *25*(1), 208-218. doi:10.1037/a0017474

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., et al. (2007). The English lexicon project. *Behavior Research Methods*, *39*(3), 445-459. doi:10.3758/BF03193014

Balota, D. A., Yap, M. J., Cortese, M. J., & Watson, J. M. (2008). Beyond mean response latency: Response time distributional analyses of semantic priming. *Journal of Memory and Language*, *59*(4), 495-523. doi:10.1016/j.jml.2007.10.004

Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition*, *10*(1), 82-93. doi:10.3758/BF03196968

Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition*, *11*(3), 211-227. doi:10.3758/BF03197629

Bates, D. M. (2005, May). Fitting linear moxed models in R. *R news*, *5*(1), 27-30.

Bates, D. M., & Sarkar, D. (2007). *lme4: Linear mixed-effects models using S4 classes*. R package.

Bayes, T., & Price, R. (1763). An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S. *Philosophical Transactions of the Royal Society of London*, *53*, 370-418. doi:10.1098/rstl.1763.0053

Biederman, G. B. (1966). Supplementary report: The recognition of tachistoscopically presented five-letter words as a function of digram frequency. *Journal of Verbal Learning and Verbal Behavior*, *5*(2), 208-209. doi:10.1016/S0022-5371(66)80020-8

Binder, J. R., McKiernan, K. A., Parsons, M. E., Westbury, C. F., Possing, E. T., Kaufman, J. N., & Buchanan, L. (2003). Neural Correlates of Lexical Access during Visual Word Recognition. *Journal of Cognitive Neuroscience*, *15*(3), 372-393. doi:10.1162/089892903321593108

Bodner, G. E., & Masson, M. E. J. (1997). Masked repetition priming of words and nonwords: Evidence for a nonlexical basis for priming. *Journal of Memory and Language*, *37*, 268-293. doi:10.1006/jmla.1996.2507

Bodner, G. E., & Masson, M. E. J. (2001). Prime Validity Affects Masked Repetition Priming: Evidence for an Episodic Resource Account of Priming. *Journal of Memory and Language*, *45*(4), 616-647. doi:10.1006/jmla.2001.2791

Bodner, G. E., & Masson, M. E. J. (2003). Beyond spreading activation: An influence of relatedness proportion om masked semantic priming. *Psychonomic Bulletin & Review*, *10*(3), 645-652. doi:10.3758/BF03196527

Bodner, G. E., & Masson, M. E. J. (2004). Beyond binary judgements: Prime validity modulates masked repetition priming in the naming task. *Memory & Cognition*, *32*(1), 1-11. doi:10.3758/BF03195815

Bodner, G. E., Masson, M. E. J., & Richard, N. T. (2006). Repetition proportion biases masked priming of lexical decisions. *Memory & Cognition*, *34*(6), 1298-1311. doi:10.3758/BF03193273

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*(4), 700-765. doi:10.1037/0033-295X.113.4.700

Bowers, J. S. (2010). Does masked and unmasked priming reflect Bayesian inference as implemented in the Bayesian Reader? *European Journal of Cognitive Psychology*, *22*(5), 779-797. doi:10.1080/09541440902955900

Bowers, J. S., Davis, C. J., & Hanley, D. A. (2005). Automatic semantic activation of embedded words: Is there a 'hat' in 'that'? *Journal of Memory and Language*, *52*(1), 131-143. doi:10.1016/j.jml.2004.09.003

Boy, F., & Sumner, P. (2010). Tight coupling between positive and reversed priming in the masked prime paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(4), 892-905. doi:10.1037/a0017173

Breitmeyer, B. G., & Hanif, W. (2008). 'Change of Mind' within and between nonconscious (masked) and conscious (unmasked) visual processing. *Consciousness and Cognition*, *17*(1), 254-266. doi:10.1016/j.concog.2007.08.001

Broadbent, D. E., & Gregory, M. H. (1968). Visual perception of words differing in letter digram frequency. *Journal of Verbal Learning and Verbal Behavior*, *7*(2), 569-571. doi:10.1016/S0022-5371(68)80052-0

Broadbent, D. E., & Gregory, M. H. (1971). Effects on tachistoscopic perception from independent variation of word probability and of letter probability. *Acta Psychologica*, *35*(1), 1-14. doi:10.1016/0001-6918(71)90027-8

Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence. *Journal of Cognitive Neuroscience*, *5*(1), 34-44. doi:10.1162/jocn.1993.5.1.34

Brown, S., Cousineau, D., & Heathcote, A. (2008). *Technical Manual for QMPE: Fortran code to fit response time distributions*.

Brown, S., & Heathcote, A. (2003). QMLE: Fast, robust, and efficient estimation of distribution functions based on quantiles. *Behavior Research Methods, Instruments, & Computers*, *35*(4), 485-492. doi:10.3758/BF03195527

Brysbaert, M. (1995). Arabic number reading: On the nature of the numerical scale and the origin of phonological recoding. *Journal of Experimental Psychology: General*, *124*, 434-452. doi:10.1037/0096-3445.124.4.434

Brysbaert, M. (2007, May 3). *'The language-as-fixed-effect fallacy': Some simple SPSS solutions to a complex problem (Version 2.0)*. Royal Holloway.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977-990. doi:10.3758/BRM.41.4.977

Burle, B., Possamaï, C.-A., Vidal, F., Bonnet, M., & Hasbroucq, T. (2002). Executive control in the Simon effect: an electromyographic and distributional analysis. *Psychological Research*, *66*, 324-336. doi:10.1007/s00426-002-0105-6

Carpenter, G., & Grossberg, S. (1987). A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics, and Image Processing*, *37*(1), 54-115. doi:10.1016/S0734-189X(87)80014-2

Chaffin, R., Morris, R. K., & Seely, R. E. (2001). Learning new word meanings from context: A study of eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(1), 225-235. doi:10.1037//0278-7393.27.1.225

Chan, E., Ybarra, O., & Schwarz, N. (2006). Reversing the affective congruency effect: The role of target word frequency of occurrence. *Journal of Experimental Social Psychology*, *42*, 365-372. doi:10.1016/j.jesp.2005.04.008

Chelouah, R., & Siarry, P. (2005). A hybrid method combining continuous tabu search and Nelder?Mead simplex algorithms for the global optimization of multiminima functions. *European Journal of Operational Research*, *161*(3), 636-654. doi:10.1016/j.ejor.2003.08.053

Cohen, J. (1977). *Statistical power analysis for the behavioural sciences* (Revised ed.). New York: Academic Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum Associates.

Colombo, L. (1986). Activation and inhibition with orthographically similar words. *Journal of Experimental Psychology: Human Perception and Performance*, *12*(2), 226-234. doi:10.1037/0096-1523.12.2.226

Coltheart, M. (2006a). Perhaps Functional Neuroimaging has not told us Anything about the Mind (So Far). *Cortex*, *42*(3), 422-427. doi:10.1016/S0010-9452(08)70374-5

Coltheart, M. (2006b). What has Functional Neuroimaging told us about the Mind (so far)? (Position Paper Presented to the European Cognitive Neuropsychology Workshop, Bressanone, 2005). *Cortex*, *42*(3), 323-331. doi:10.1016/S0010-9452(08)70358-7

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, *100*(4), 589-608. doi:10.1037/0033-295X.100.4.589

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornič (Ed.), *Attention and Performance IV: Proceedings of the Sixth International Symposium on Attention and Performance, Stockholm, Sweden, July 28-August 1, 1975* (pp. 535-555).

Coltheart, M., & Rastle, K. (1994). Serial processing in reading aloud: Evidence for dual-route models of reading. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(6), 1197-1211. doi:10.1037/0096-1523.20.6.1197

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204-256. doi:10.1037//0033-295X.108.1.204

Cortese, M. J., & Khanna, M. M. (2007). Age of acquisition predicts naming and lexical-decision performance above and beyond 22 other predictor variables: An analysis of 2,342 words. *The Quarterly Journal of Experimental Psychology*, *60*(8), 1072-1082. doi:10.1080/17470210701315467

Cousineau, D., Brown, S., & Heathcote, A. (2004). Fitting distributions using maximum likelihood: Methods and packages. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 742-756. doi:10.3758/BF03206555

Damerau, F. J. (1964). A technique for computer detection and correction of spelling errors. *Communications of the ACM*, *7*(3), 171-176. doi:10.1145/363958.363994

Damian, M. F. (2001). Congruity effects evoked by subliminally presented primes: Automaticity rather than semantic processing. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 154-165. doi:10.1037/0096-1523.27.1.154

Davis, C. J. (1999). *The Self-Organising Lexical Acquisition and Recognition (SOLAR) model of visual word recognition*. University of New South Wales.

Davis, C. J. (2003). Factors underlying masked priming effects in competitive network models of visual word recognition. In S. Kinoshita & S. J. Lupker (Eds.), *Masked priming: The state of the art*, Macquarie Monographs in Cognitive Science (pp. 67-94). New York and Hove: Psychology Press.

Davis, C. J. (2005). N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods*, *37*, 65-70. doi:10.3758/BF03206399

Davis, C. J. (2010). The spatial coding model. *Psychological Review*, *117*(3), 713-758. doi:10.1037/a0019738

Davis, C. J., & Bowers, J. S. (2004). What do letter migration errors reveal about letter position coding in visual word recognition? *Journal of Experimental Psychology: Human Perception and Performance*, *30*(5), 923-941. doi:10.1037/0096-1523.30.5.923

Davis, C. J., & Bowers, J. S. (2006). Contrasting five different theories of letter position coding: Evidence from orthographic similarity effects. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(3), 535-557. doi:10.1037/0096-1523.32.3.535

Davis, C. J., & Lupker, S. J. (2006). Masked inhibitory priming in English: Evidence for lexical inhibition. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(3), 668-687. doi:10.1037/0096-1523.32.3.668

Davis, C. J., Perea, M., & Acha, J. (2009). Re(de)fining the orthographic neighborhood: The role of addition and deletion neighbors in lexical decision and reading. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(5), 1550-1570. doi:10.1037/a0014253

Davis, C. J., & Taft, M. (2005). More words in the neighborhood: Interference in lexical decision due to deletion neighbors. *Psychonomic Bulletin & Review*, *12*(5), 904-910. doi:10.3758/BF03196784

De Bleser, R., Cholewa, J., Stadie, N., & Tabatabaie, S. (1997). LeMo, an Expert System for Single Case Assessment of W ord Processing Impairments in Aphasic Patients. *Neuropsychological Rehabilitation*, *7*(4), 339-366. doi:10.1080/713755540

Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J.-F., Poline, J.-B., & Rivière, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience*, *4*(7), 752-758. doi:10.1038/89551

Dehaene, S., Naccache, L., Le Clec'H, G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., Van de Moortele, P.-F., et al. (1998). Imaging unconscious semantic priming. *Nature*, *395*(6702), 597–599. doi:10.1038/26967

Ditterich, J., Mazurek, M. E., & Shadlen, M. N. (2003). Microstimulation of visual cortex affects the speed of perceptual decisions. *Nature Neuroscience*, *6*(8), 891-898. doi:10.1038/nn1094

Dorfman, D., & Glanzer, M. (1988). List composition effects in lexical decision and recognition memory. *Journal of Memory and Language*, *27*, 633-648. doi:10.1016/0749-596X%2888%2990012-5

Duyck, W., Desmet, T., Verbeke, L. P. C., & Brysbaert, M. (2004). WordGen: A tool for word selection and nonword generationin Dutch, English, German, and French. *Behavior Research Methods, Instruments, & Computers*, *36*, 488-499. doi:10.3758/BF03195595

Eimer, M., & Schlaghecken, F. (2003). Response facilitation and inhibition in subliminal priming. *Biological Psychology*, *64*(1-2), 7-26. doi:10.1016/S0301-0511(03)00100-5

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A Dynamical Model of Saccade Generation During Reading. *Psychological Review*, *112*(4), 777-813. doi:10.1037/0033-295X.112.4.777

Evett, L. J., & Humphreys, G. W. (1981). The use of abstract graphemic information in lexical access. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *33A*(4), 325-350. doi:10.1080/14640748108400797

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175-191. doi:10.3758/BF03193146

Felici, A., & Pal, P. (2008). A probabilistic approach to language structure (pp. 28-31). Presented at the Quantitative Investigations in Theoretical Linguistics (QUITL3),

Helsinki, Suomi Finland. Retrieved from http://www.ling.helsinki.fi/sky/tapahtumat/qitl/QITL3_Proceedings.pdf

Ferrand, L., & Grainger, J. (1992). Phonology and orthography in visual word recognition: Evidence from masked non-word priming. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *45A*(3), 353-372. doi:10.1080/02724989208250619

Ferrand, L., & Grainger, J. (1994). Effects of orthography are independent of phonology in masked form priming. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *47A*(3), 365-382. doi:10.1080/02724989208250619

Ferrand, L., New, B., Brysbaert, M., Keulers, E., Bonin, P., Méot, A., Augustinova, M., et al. (2010). The French Lexicon Project: Lexical decision data for 38,840 French words and 38,840 pseudowords. *Behavior Research Methods*, *42*(2), 488-496. doi:10.3758/BRM.42.2.488

Fischler, I., & Bloom, P. A. (1979). Automatic and attentional processes in the effects of sentence contexts on word recognition. *Journal of Verbal Learning and Verbal Behavior*, *18*(1), 1-20. doi:10.1016/S0022-5371(79)90534-6

Forster, K. I. (1985). Lexical acquisition and the modular lexicon. *Language and Cognitive Processes*, *1*(2), 87-108. doi:10.1080/01690968508402073

Forster, K. I. (1998). The Pros and Cons of Masked Priming. *Journal of Psycholinguistic Research*, *27*(2), 203-233. doi:10.1023/A:1023202116609

Forster, K. I. (1999). The Microgenesis of Priming Effects in Lexical Access. *Brain and Language*, *68*, 5-15. doi:10.1006/brln.1999.2078

Forster, K. I. (2004). Category size effects revisited: Frequency and masked priming effects in semantic categorization. *Brain and Language*, *90*, 276-286. doi:10.1016/S0093-934X%2803%2900440-1

Forster, K. I. (2009). The intervenor effect in masked priming: How does masked priming survive across an intervening word? *Journal of Memory and Language*, *60*(1), 36-49. doi:10.1016/j.jml.2008.08.006

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 680-698. doi:10.1037/0278-7393.10.4.680

Forster, K. I., & Davis, C. (1991). The density constraint on form-priming in the naming task: Interference effects from a masked prime. *Journal of Memory and Language*, *30*, 1-25. doi:10.1016/0749-596X%2891%2990008-8

Forster, K. I., Davis, C., Schoknecht, C., & Carter, R. (1987). Masked priming with graphemically related forms: Repetition or partial activation? *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *39A*(2), 211-251. doi:10.1080/14640748708401785

Forster, K. I., & Forster. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*, 116-124. doi:10.3758/BF03195503

Forster, K. I., Mohan, K., & Hector, J. (2003). The mechanics of masked priming. In S. Kinoshita & S. J. Lupker (Eds.), *Masked priming: The state of the art*, Macquarie Monographs in Cognitive Science (pp. 2-21). New York and Hove: Psychology Press.

Frost, R., Forster, K. I., & Deutsch, A. (1997). What can we learn from the morphology of Hebrew? A masked-priming investigation of morphological representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(4), 829-856. doi:10.1037/0278-7393.23.4.829

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, *113*(2), 256-281. doi:10.1037/0096-3445.113.2.256

Glover, F. (1989). Tabu Search--Part I. *INFORMS Journal on Computing*, *1*(3), 190-206. doi:10.1287/ijoc.1.3.190

Glover, F. (1990). Tabu Search--Part II. *INFORMS Journal on Computing*, *2*(1), 4-32. doi:10.1287/ijoc.2.1.4

Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language*, *29*(2), 228-244. doi:10.1016/0749-596X(90)90074-A

Grainger, J. (2008). Cracking the orthographic code: An introduction. *Language and Cognitive Processes*, *23*(1), 1-35. doi:10.1080/01690960701578013

Grainger, J., Colé, P., & Seguí, J. (1991). Masked morphological priming in visual word recognition. *Journal of Memory and Language*, *30*(3), 370-384. doi:10.1016/0749-596X%2891%2990042-I

Grainger, J., & Ferrand, L. (1996). Masked Orthographic and Phonological Priming in Visual Word Recognition and Naming: Cross-Task Comparisons. *Journal of Memory and Language*, *35*(5), 623-647. doi:10.1006/jmla.1996.0033

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: a multiple read-out model. *Psychological Review*, *103*(3), 518-565. doi:10.1037/0033-295X.103.3.518

Grainger, J., & Jacobs, A. M. (1999). Temporal integration of information in orthographic priming. *Visual Cognition*, *6*(3), 461-492. doi:10.1080/135062899395064

Grainger, J., Kiyonaga, K., & Holcomb, P. J. (2006). The Time Course of Orthographic and Phonological Code Activation. *Psychological Science*, *17*(12), 1021-1026. doi:10.1111/j.1467-9280.2006.01821.x

Grainger, J., O'Regan, J. K., Jacobs, A. M., & Seguí, J. (1989). On the role of competing word units in visual word recognition: The neighbourhood frequency effect. *Perception & Psychophysics*, *45*(3), 189-195. doi:10.3758/BF03210696

Grainger, J., O'Regan, J. K., Jacobs, A. M., & Seguí, J. (1992). Neighborhood frequency effects and letter visibility in visual word recognition. *Perception & Psychophysics*, *51*(1), 49-56. doi:10.3758/BF03205073

Greenwald, A. G., Abrams, R. L., Naccache, L., & Dehaene, S. (2003). Long-term semantic memory versus contextual memory in unconscious number

processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 235-247. doi:10.1037/0278-7393.29.2.235

Greenwald, A. G., Draine, S. C., & Abrams, R. L. (1996). Three cognitive markers of unconscious semantic activation. *Science*, *273*, 1699-1702. doi:10.1126/science.273.5282.1699

Hall, A. T. (2000). *Phonologie : eine Einführung*. Berlin: De Gruyter.

Hamming, R. W. (1950). Error detecting and error correcting codes. *The Bell System Technical Journal*, *29*(2), 147-160.

Hauk, O., Patterson, K., Woollams, A. M., Watling, L., Pulvermuller, F., & Rogers, T. T. (2006). [Q:] When Would You Prefer a SOSSAGE to a SAUSAGE? [A:] At about 100 msec. ERP Correlates of Orthographic Typicality and Lexicality in Written Word Recognition. *Journal of Cognitive Neuroscience*, *18*, 818-832. doi:10.1162/jocn.2006.18.5.818

Heathcote, A., Brown, S., & Cousineau, D. (2004). QMPE: Estimating Lognormal, Wald, and Weibull RT distributions with a paramter-dependent lower bound. *Behavior Research Methods, Instruments, & Computers*, *36*(2), 277-290. doi:10.3758/BF03195574

Heyer, G., Quasthoff, U., & Wittig, T. (2006). *Text Mining: Wissensrohstoff Text : Konzepte, Algorithmen, Ergebnisse*. Herdecke [u.a.]: W3L-Verlag.

Howes, D. H., & Solomon, R. L. (1951). Visual duration threshold as a function of word-probability. *Journal of Experimental Psychology*, *41*(6), 401-410. doi:10.1037/h0056020

Humphreys, G. W., Besner, D., & Quinlan, P. T. (1988). Event perception and the word repetition effect. *Journal of Experimental Psychology: General*, *117*(1), 51-67. doi:10.1037/0096-3445.117.1.51

Humphreys, G. W., Quinlan, P. T., Evett, L. J., & Besner, D. (1987). Orthographic priming: Qualitative differences between priming from identified and unidentified primes. In M. Coltheart (Ed.), *Attention and Performance XII: The Psychology of*

Reading (pp. 106-125). Hove and London (UK), Hillsdale (USA): Lawrence Erlbaum Associates.

Humphreys, J. (2008). *Wirds, werds, wyrdz: Visual wordlikeness, lexical phonology, and models of visual word recognition*. University of Bristol, Bristol.

Inhoff, A. W. (1984). Two stages of word processing during eye fixations in the reading of prose. *Journal of Verbal Learning and Verbal Behavior*, *23*(5), 612-624. doi:10.1016/S0022-5371(84)90382-7

International Phonetic Association. (1999). *Handbook of the International Phonetic Association : a guide to the use of the International Phonetic Alphabet*. Cambridge U.K. ;;New York NY: Cambridge University Press.

Irwin, D. E. (1998). Lexical Processing during Saccadic Eye Movements. *Cognitive Psychology*, *36*(1), 1-27. doi:10.1006/cogp.1998.0682

Jacobs, A. M., & Grainger, J. (1992). Testing a semistocastic variant of the interactive activation model in different word recognition experiments. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(4), 1174-1188. doi:10.1037/0096-1523.18.4.1174

Jacobs, A. M., Grainger, J., & Ferrand, L. (1995). The incremental priming technique: A method for determining within-condition priming effects. *Perception & Psychophysics*, *57*, 1101-1110. doi:10.3758/BF03208367

James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance*, *104*(2), 130-136. doi:10.1037/0096-1523.1.2.130

Jared, D. (1997). Spelling–sound consistency affects the naming of high-frequency Words. *Journal of Memory and Language*, *36*(4), 505-529. doi:10.1006/jmla.1997.2496

Jastrzembski, J. E. (1981). Multiple meanings, number of related meanings, frequency of occurrence, and the lexicon. *Cognitive Psychology*, *13*(2), 278-305. doi:10.1016/0010-0285(81)90011-6

Java Runtime Environment. (2011). 500 Oracle Parkway, Redwood Shores, California USA: Oracle Corporation. Retrieved from http://www.java.com

Kahneman, D. (1968). Method, findings, and theory in studies of visual masking. *Psychological Bulletin*, *70*(6), 404-425. doi:10.1037/h0026731

Kay, J., Lesser, R., & Coltheart, M. (1996). Psycholinguistic assessments of language processing in aphasia (PALPA): An introduction. *Aphasiology*, *10*(2), 159-180. doi:10.1080/02687039608248403

Kennison, S. M., & Clifton, C. (1995). Determinants of parafoveal preview benefit in high and low working memory capacity readers: Implications for eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(1), 68-81. doi:10.1037/0278-7393.21.1.68

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, *42*(3), 627-633. doi:10.3758/BRM.42.3.627

Kiesel, A., Berner, M. P., & Kunde, W. (2008). Negative congruency effects: A test of the inhibition account. *Consciousness and Cognition*, *17*(1), 1-21. doi:10.1016/j.concog.2006.11.003

Kiesel, A., Kunde, W., Pohl, C., Berner, M. P., & Hoffmann, J. (2009). Playing Chess Unconsciously. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 292-298. doi:10.1037/a0014499

Kinoshita, S., & Hunt, L. (2008). RT distribution analysis of category congruence effects with masked primes. *Memory & Cognition*, *36*, 1324-1334. doi:10.3758/MC.36.7.1324

Kinoshita, S., & Lupker, S. J. (Eds.). (2003). *Masked priming: The state of the art*. Macquarie Monographs in Cognitive Science. New York and Hove: Psychology Press.

Kinoshita, S., & Norris, D. (2010). Masked priming effect reflects evidence accumulated by the prime. *The Quarterly Journal of Experimental Psychology*, *63*(10), 194-204. doi:10.1080/17470210902957174

Klauer, C. K., Eder, A. B., Greenwald, A. G., & Abrams, R. L. (2007). Priming of semantic classifications by novel subliminal prime words. *Consciousness and Cognition*, *16*, 63-83. doi:10.1016/j.concog.2005.12.002

Klinger, M. R., Burton, P. C., & Pitts, G. S. (2000). Mechanisms of unconscious priming: I. Response competition, not spreading activation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(2), 441-455. doi:10.1037//0278-7393.26.2.441

Kousta, S.-T., Vinson, D. P., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition*, *112*(3), 473-481. doi:10.1016/j.cognition.2009.06.007

Krifka, M. (1992). Thematic relations as links between nominal reference and temporal constitution. In A. Szabolcsi & I. A. Sag (Eds.), *Lexical matters* (pp. 29-54). Stanford CA: Center for the Study of Language and Information.

Kroll, J. F., & Merves, J. S. (1986). Lexical access for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(1), 92-107. doi:10.1037/0278-7393.12.1.92

Kučera, H., & Francis, W. (1967). *Computational analysis of present-day American English*. RI: Brown University Press.

Kunde, W., Kiesel, A., & Hoffmann, J. (2003). Conscious control over the content of unconscious cognition. *Cognition*, *88*, 223-242. doi:10.1016/S0010-0277%2803%2900023-4

Küpfmüller, K. (1949). *Die Systemtheorie der elektrischen Nachrichtenübertragung* (erste Auflage.). Stuttgart: Hirzel Verlag.

Ladefoged, P., & Johnston, K. (2011). *A course in phonetics* (6th ed.). Boston MA: Wadsworth Cenage.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*(2), 211-240. doi:10.1037/0033-295X.104.2.211

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physics Doklady*, *10*(8), 707-710.

Levy-Drori, S., & Henik, A. (2006). Concreteness and context availability in lexical decision tasks. *The American Journal of Psychology*, *119*(1), 45-65.

Loth, S. (2007, unpublished). *Words and nonwords*. Universität Leipzig, Leipzig.

Lukatela, G., Frost, S. J., & Turvey, M. T. (1998). Phonological Priming by Masked Nonword Primes in the Lexical Decision Task. *Journal of Memory and Language*, *39*(4), 666-683. doi:10.1006/jmla.1998.2599

Lupker, S. J., Perea, M., & Davis, C. J. (2008). Transposed-letter effects: Consonants, vowels and letter frequency. *Language and Cognitive Processes*, *23*(1), 93-116. doi:10.1080/01690960701579714

Manning, C. D., & Schütze, H. (1999). *Foundations of statistical natural language processing*. Cambridge Mass.: MIT Press.

Massaro, D. W., Venezky, R. L., & Taylor, G. A. (1979). Orthographic regularity, positional frequency, and visual processing of letter strings. *Journal of Experimental Psychology: General*, *108*(1), 107-124. doi:10.1037/0096-3445.108.1.107

Mayzner, M. S., & Tresselt, M. E. (1962a). Anagram solution times: A function of word transition probabilities. *Journal of Experimental Psychology*, *63*(5), 510-513. doi:10.1037/h0043721

Mayzner, M. S., & Tresselt, M. E. (1962b). The ranking of letter pairs and single letters to match digram and single-letter frequency counts. *Journal of Verbal Learning and Verbal Behavior*, *1*(3), 203-207. doi:10.1016/S0022-5371(62)80032-2

Mayzner, M. S., & Tresselt, M. E. (1963). Anagram solution times: A function of word length and letter position variables. *Journal of Psychology: Interdisciplinary and Applied*, *55*(2), 469-475. doi:10.1080/00223980.1963.9916641

Mayzner, M. S., & Tresselt, M. E. (1966). Anagram solution times: A function of multiple-solution anagrams. *Journal of Experimental Psychology*, *71*(1), 66-73. doi:10.1037/h0022665

McClelland, J. L., & Johnston, J. C. (1977). The role of familiar units in perception of words and nonwords. *Perception & Psychophysics*, *22*(3), 249-261. doi:10.3758/BF03199687

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, *88*(5), 375-407. doi:10.1037/0033-295X.88.5.375

McCormick, S. F., Brysbaert, M., & Rastle, K. (2009). Is morphological decomposition limited to low-frequency words? *The Quarterly Journal of Experimental Psychology*, *62*(9), 1706-1715. doi:10.1080/17470210902849991

Merikle, P. M., Coltheart, M., & Lowe, D. G. (1971). On the selective effects of a patterned masking stimulus. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *25*(3), 264-279. doi:10.1037/h0082388

Mewhort, D. J. K., & Campbell, A. J. (1978). Processing spatial information and the selective-masking effect. *Perception & Psychophysics*, *24*(1), 93-101. doi:10.3758/BF03202978

Miller, G. A., Bruner, J. S., & Postman, L. (1954). Familiarity of letter sequences and tachistoscopic identification. *Journal of General Psychology*, *50*, 129-139.

Morrison, C. M., & Ellis, A. W. (2000). Real age of acquisition in word naming and lexicl decision. *British Journal of Psychology*, *91*, 167-180. doi:10.1348/000712600161763

Murray, W. S., & Forster, K. I. (2004). Serial mechanisms in lexical access: The rank hypothesis. *Psychological Review*, *111*(3), 721-756. doi:10.1037/0033-295X.111.3.721

Murray, W. S., & Forster, K. I. (2008). The rank hypothesis and lexical decision: A reply to Adelman and Brown (2008). *Psychological Review*, *115*(1), 240-251. doi:10.1037/0033-295X.115.1.240

Naccache, L., & Dehaene, S. (2001). Unconscious semantic priming extends to novel unseen stimuli. *Cognition*, *80*, 215-229. doi:10.1016/S0010-0277%2800%2900139-6

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *The Computer Journal*, *7*(4), 308-313. doi:10.1093/comjnl/7.4.308

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 516-524. doi:10.3758/BF03195598

Norris, D. (2006). The Bayesian reader: Explaining word recognition as an optimal Bayesian decision process. *Psychological Review*, *113*(2), 327-357. doi:10.1037/0033-295X.113.2.327

Norris, D., & Kinoshita, S. (2008). Perception as Evidence Accumulation and Bayesian Inference: Insights From Masked Priming. *Journal of Experimental Psychology: General*, *137*(3), 434-455. doi:10.1037/a0012799

Owsowitz, S. E. (1963). The effects of word familiarity and letter structure familiarity on the perception of words. *RAND publication*. Retrieved from http://www.rand.org/pubs/papers/P2820

Oxford Journals Reports — Most-Cited Articles as of August 1, 2011. (2011, August).*Oxford University Press - Reports*. Retrieved June 2, 2011, from http://comjnl.oxfordjournals.org/reports/most-cited

Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W. (1982). An activation-verification model for letter and word recognition: The word-superiority effect. *Psychological Review*, *89*(5), 573-594. doi:10.1037/0033-295X.89.5.573

Page, M. P. A. (2006). What Can't Functional Neuroimaging Tell the Cognitive Psychologist? *Cortex*, *42*(3), 428-443. doi:10.1016/S0010-9452(08)70375-7

Perea, M., Fernández, L., & Rosa, E. (1998). El papel del status léxico y de la frecuencia del estímulo-señal en la condición no relacionada con la técnica de presentación enmascarada del estímulo-señal. *Psicológica*, *19*, 311-319.

Perea, M., Gómez, P., & Fraga, I. (2010). Masked nonword repetition effects in yes/no and go/no-go lexical decision: A test of the evidence accumulation and deadline

accounts. *Psychonomic Bulletin & Review*, *17*(3), 369-374. doi:10.3758/PBR.17.3.369

Perea, M., & Gotor, A. (1997). Associative and semantic priming effects occur at very short stimulus-onset asynchronies in lexical decision and naming. *Cognition*, *62*, 223-240. doi:10.1016/S0010-0277%2896%2900782-2

Perea, M., & Lupker, S. J. (2003a). Transposed-letter confusability effects in masked form priming. In S. Kinoshita & S. J. Lupker (Eds.), *Masked priming: The state of the art* (pp. 97-120). New York, NY, US: Psychology Press.

Perea, M., & Lupker, S. J. (2003b). Does jugde activate COURT? Transposed-letter similarity in masked associative priming. *Memory & Cognition*, *31*(6), 829-841. doi:10.3758/BF03196438

Pollatsek, A., & Well, A. D. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(3), 785-794. doi:10.1037/0278-7393.21.3.785

Pompino-Marschall, B. (2009). *Einführung in die Phonetik* (3., durchges. Aufl.). Berlin, New York NY: De Gruyter.

Postman, L., & Schneider, B. H. (1951). Personal values, visual recognition, and recall. *Psychological Review*, *58*(4), 271-284. doi:10.1037/h0062787

Quinn, W. M., & Kinoshita, S. (2008). Congruence effect in semantic categorization with masked primes with narrow and broad categories. *Journal of Memory and Language*, *58*, 286-306. doi:10.1016/j.jml.2007.03.004

R development core team. (2007). *R: A language and environment for statistical computing*. Wien, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Raaijmakers, J. G. W., Schrijnmakers, J. M. C., & Gremmen, F. (1999). How to deal with 'The language-as-fixed-effect fallacy': Common misconceptions and alternative solutions. *Journal of Memory and Language*, *41*(3), 416-426. doi:10.1006/jmla.1999.2650

Ramsden, N., & Ramsden, L. (2007). *Word Building and Spelling*. Retrieved from http://www.neilramsden.co.uk/spelling/

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59-108. doi:10.1037/0033-295X.85.2.59

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin*, *86*(3), 446-461. doi:10.1037/0033-2909.86.3.446

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A Diffusion Model Account of the Lexical Decision Task. *Psychological Review*, *111*(1), 159-182. doi:10.1037/0033-295X.111.1.159

Ratcliff, R., & Murdock, B. B. (1976). Retrieval processes in recognition memory. *Psychological Review*, *83*(3), 190-214. doi:10.1037/0033-295X.83.3.190

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372-422. doi:10.1037/0033-2909.124.3.372

Rayner, K. (2009). Eye movements in reading: Models and data. *Journal of Eye Movement Research*, *2*(5), 1-10.

Reynvoet, B., Gevers, W., & Caessens, B. (2005). Unconscious Primes Activate Motor Codes Through Semantics. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 991-1000. doi:10.1037/0278-7393.31.5.991

Rice, G. A., & Robinson, D. O. (1975). The role of bigram frequency in the perception of words and nonwords. *Memory & Cognition*, *3*(5), 513-518. doi:10.3758/BF03197523

Richter, E. M., Engbert, R., & Kliegl, R. (2006). Current advances in SWIFT. *Cognitive Systems Research*, *7*(1), 23-33. doi:10.1016/j.cogsys.2005.07.003

Rouder, J. N., & Speckman, P. L. (2004). An evaluation of the vincentizing method of forming group-level response time distributions. *Psychonomic Bulletin & Review*, *11*(3), 419-427. doi:10.3758/BF03196589

Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior*, *9*(5), 487-494. doi:10.1016/S0022-5371(70)80091-3

Rubenstein, H., Lewis, S. S., & Rubenstein, M. A. (1971). Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior*, *10*(6), 645-657. doi:10.1016/S0022-5371(71)80071-3

Rubenstein, H., Richter, M. L., & Kay, E. J. (1975). Pronounceability and the visual recognition of nonsense words. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 651-657. doi:10.1016/S0022-5371(75)80053-3

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: II. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review*, *89*(1), 60-94. doi:10.1037/0033-295X.89.1.60

Rumelhart, D. E., & Siple, P. (1974). Process of recognizing tachistoscopically presented words. *Psychological Review*, *81*(2), 99-118. doi:10.1037/h0036117

Schlaghecken, F., Bowman, H., & Eimer, M. (2006). Dissociating Local and Global Levels of Perceptuo-Motor Control in Masked Priming. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(3), 618-632. doi:10.1037/0096-1523.32.3.618

Schoonbaert, S., & Grainger, J. (2004). Letter position coding in printed word perception: Effects of repeated and transposed letters. *Language and Cognitive Processes*, *19*(3), 333-367. doi:10.1080/01690960344000198

Schuberth, R. E., & Eimas, P. D. (1977). Effects of context on the classification of words and nonwords. *Journal of Experimental Psychology: Human Perception and Performance*, *3*(1), 27-36. doi:10.1037/0096-1523.3.1.27

Schwanenflugel, P. J., Kipp Harnishfeger, K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words*1. *Journal of Memory and Language*, *27*(5), 499-520. doi:10.1016/0749-596X(88)90022-8

Sears, C. R., Hino, Y., & Lupker, S. J. (1995). Neighborhood size and neighborhood frequency effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(4), 876-900. doi:10.1037/0096-1523.21.4.876

Seguí, J., & Grainger, J. (1990). Priming word recognition with orthographic neighbors: Effects of relative prime-target frequency. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(1), 65-76. doi:10.1037/0096-1523.16.1.65

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*(4), 523-568. doi:10.1037/0033-295X.96.4.523

Sereno, J. A. (1991). Graphemic, associative, and syntactic priming effects at a brief stimulus onset asynchrony in lexical decision and naming. *Journal*, *17*, 459-477. doi:10.1037/0278-7393.17.3.459

Shadlen, M. N., & Newsome, W. T. (2001). Neural Basis of a Perceptual Decision in the Parietal Cortex (Area LIP) of the Rhesus Monkey. *Journal of Neurophysiology*, *86*(4), 1916 -1936.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *5*(1), 3. doi:10.1145/584091.584093

Shannon, C. E. (1951). Prediction and entropy of printed English. *Bell System Technical Journal*, *30*(1), 50-64.

Siakaluk, P. D., Sears, C. R., & Lupker, S. J. (2002). Orthographic neighborhood effects in lexical decision: The effects of nonword orthographic neighborhood size. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(3), 661-681. doi:10.1037/0096-1523.28.3.661

Siegel, P., & Weinberger, J. (2009). Very brief exposure: The effects of unreportable stimuli on fearful behavior. *Consciousness and Cognition*, *18*, 939-951. doi:10.1016/j.concog.2009.08.001

Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences*, *27*(3), 161-168. doi:10.1016/j.tins.2004.01.006

Smith, P. L., Ratcliff, R., & Wolfgang, B. J. (2004). Attention orienting and the time course of perceptual decisions: response time distributions with masked and unmasked displays. *Vision Research*, *44*(12), 1297-1320. doi:10.1016/j.visres.2004.01.002

Solomon, R. L., & Howes, D. H. (1951). Word frequency, personal values, and visual duration thresholds. *Psychological Review*, *58*(4), 256-270. doi:10.1037/h0058228

Stroop, J. R. (1992). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology: General*, *121*, 15-23. doi:10.1037/0096-3445.121.1.15

The British National Corpus. (2001). Oxford: Oxford University Computing Services on the bhalf of the BNC Consortium.

Tse, C.-S., Balota, D. A., Yap, M. J., Duchek, J. M., & McCabe, D. P. (2010). Effects of healthy aging and early stage dementia of the Alzheimer's type on components of response time distributions in three attention tasks. *Neuropsychology*, *24*(3), 300-315. doi:10.1037/a0018274

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*(3), 550-592. doi:10.1037//0033-295X.108.3.550

Van Assche, E., & Grainger, J. (2006). A Study of Relative-Position Priming With Superset Primes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(2), 399-415. doi:10.1037/0278-7393.32.2.399

van Zandt, T. (2000). How to fit a response time distribution. *Psychonomic Bulletin & Review*, *7*(3), 424-465. doi:10.3758/BF03214357

Vickers, D. (1970). Evidence for an Accumulator Model of Psychophysical Discrimination. *Ergonomics*, *13*(1), 37-58. doi:10.1080/00140137008931117

Vincent, S. B. (1912). *The function of the vibrissae in the behavior of the white rat*. University of Chicago, Chicago. Retrieved from Univeristy of Chicago.

Wang, X.-J. (2002). Probabilistic Decision Making by Slow Reverberation in Cortical Circuits. *Neuron*, *36*(5), 955-968. doi:10.1016/S0896-6273(02)01092-9

Web of Knowledge - Citation Report. (2011).*Web of Knowledge*. Retrieved August 30, 2011, from http://apps.webofknowledge.com/OutboundService.do?action=go

Westbury, C. F., & Buchanan, L. (1999). What's in a word? A sublexical effect in a lexical decision task. In M. Hahn & C. Stoness (Eds.), *Proceedings of the Twenty First Annual Conference of the Cognitive Science Society* (p. 824). Mahwah, N.J.: Lawrence Erlbaum.

Westbury, C. F., & Buchanan, L. (2002). The probability of the least likely non-length-controlled bigram affects lexical decision reaction times. *Brain and Language*, *81*(1-3), 66-78. doi:10.1006/brln.2001.2507

Westbury, C. F., Buchanan, L., Sanderson, M., Rhemtulla, M., & Phillips, L. (2003). Using genetic programming to discover nonlinear variable interactions. *Behavior Research Methods, Instruments, & Computers*, *35*(2), 202-216. doi:10.3758/BF03202543

Westbury, C. F., Hollis, G., & Shaoul, C. (2007). LINGUA: The language-independent neighobourhood generator of the University of Alberta. *The Mental Lexicon*, *2*(2), 271-286.

Whaley, C. P. (1978). Word-Nonword Classification Time. *Journal of Verbal Learning and Verbal Behavior*, *17*, 143-154. doi:10.1016/S0022-5371%2878%2990110-X

White, S. J., Johnson, R. L., Liversedge, S. P., & Rayner, K. (2008). Eye movements when reading transposed text: The importance of word-beginning letters. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(5), 1261-1276. doi:10.1037/0096-1523.34.5.1261

Wittgenstein, L. (1989). *Tractatus logico-philosophicus. Tagebücher 1914-1916. Philosophische Untersuchungen*. Werkausgabe in 8 Bänden (5. Auflage.). Frankfurt am Main: Suhrkamp Verlag.

Woollams, A. M., Silani, G., Okada, K., Patterson, K., & Price, C. J. (2011). Word or word-like? Dissociating orthographic typicality from lexicality in the left occipito-

temporal cortex. *Journal of Cognitive Neuroscience*, *23*(4), 992-1002. doi:10.1162/jocn.2010.21502

Yap, M. J., Balota, D. A., Cortese, M. J., & Watson, J. M. (2006). Single- versus dual-process models of lexical decision performance: Insights from response time distributional analysis. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(6), 1324-1344. doi:10.1037/0096-1523.32.6.1324

Yap, M. J., Balota, D. A., Tse, C.-S., & Besner, D. (2008). On the additive effects of stimulus quality and word frequency in lexical decision: Evidence for opposing interactive influences revealed by RT distributional analyses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(3), 495-513. doi:10.1037/0278-7393.34.3.495

Yarkoni, T., Balota, D. A., & Yap, M. J. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review*, *15*(5), 971-979. doi:10.3758/PBR.15.5.971

Yatabe, K., Pickering, M. J., & McDonald, S. A. (2009). Lexical processing during saccades in text comprehension. *Psychonomic Bulletin & Review*, *16*(1), 62-66. doi:10.3758/PBR.16.1.62

Zandt, T., Colonius, H., & Proctor, R. W. (2000). A comparison of two response time models applied to perceptual matching. *Psychonomic Bulletin & Review*, *7*(2), 208-256. doi:10.3758/BF03212980

Zorzi, M., Houghton, G., & Butterworth, B. (1998). Two routes or one in reading aloud? A connectionist dual-process model. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(4), 1131-1161. doi:10.1037/0096-1523.24.4.1131

Appendix A: Analysis of prime novelty and adaptation in Experiment 2

As briefly outlined in 3.4.1 the primes and targets in Experiment 2 were drawn from the same pool of items. This means that an effect of prime novelty could have emerged (see Damian, 2001). In the following I present the analysis of prime novelty. Additionally, participants could adapted to the task during the experimental session which could have influenced the priming effect as well. The respective analysis is presented as well.

A.1 Prime novelty

If the congruency priming effect emerged as a result of stimulus-response mappings only used primes, i.e. primes that were presented as targets already, would have elicited an effect. This contrasts with novel primes where no such memory trace could have been established and thus, no priming effect is predicted. If the stimulus-response mapping account can explain the findings an interaction of prime novelty and congruency is expected.

Analysis

Each trial of Experiment 2 was tagged with regards to prime novelty and the data were used in a post-hoc analysis. The reaction times and error scores are shown in Table A.1 for word and in Table A.2 for nonword targets (see Table 3.1 for results ignoring prime novelty). The respective ANOVAs for word and nonword targets were computed using the *z*-scores of reaction time as a dependent variable. Prime novelty and congruency were treated as a repeated factors and list as a random factor.

Results

The analysis showed a main effect of congruency like in previous analyses. There was no significant main effect of prime novelty in words by participants [$F_1(1, 36) = 1.640$, $p=0.208$], but the effect was significant by items [$F_2(1, 98) = 6.355$, $p=0.013$, $f=0.25$] indicating faster responses in trials with used than in trials with novel primes. In nonword targets the main effect of prime novelty was significant by participants [$F_1(1, 36) = 7.296$, $p=0.010$, $f=0.45$] indicating faster responses in trials with used than in trials with novel primes. But this effect was not significant by items [$F_2(1, 98) = 2.663$, $p=0.106$]. The interaction between congruency and prime novelty was not significant in word targets [$F_1(1, 36) = 0.349$, $p=0.559$; $F_2(1, 97) = 0.743$, $p=0.391$]. There was a tendency in nonword targets [$F_1(1, 36) = 3.492$, $p=0.070$; $F_2(1, 98) = 2.445$, $p=0.121$] indicating that the congruency effect in used primes was larger than in novel primes.

Table A.1: Mean reaction times in ms and error ratio in percent in Word targets of Experiment 2 as a function of congruency and prime novelty.

	Prime novelty			
	Novel		Used	
Congruency	RT	Error	RT	Error
Incongruent	501	2.99	496	4.66
Congruent	476	2.11	475	2.15
Effect	25	0.77	21	2.51

Table A.2: Mean reaction times in ms and error ratio in percent in Nonword targets of Experiment 2 as a function of congruency and prime novelty.

	Prime novelty			
	Novel		Used	
Congruency	RT	Error	RT	Error
Incongruent	523	2.78	516	4.76
Congruent	503	1.94	492	1.85
Effect	20	0.84	24	2.91

The analysis of error scores showed no significant main effect of prime novelty in word targets [$F_1(1, 36) = 2.284$, $p=0.139$; $F_2(1, 97) = 2.210$, $p=0.140$] or nonword targets [$F_1(1, 36) = 2.078$, $p=0.158$; $F_2(1, 98) = 2.662$, $p=0.106$]. There was a tendency for an interaction of congruency and prime novelty but this was not significant in word targets [$F_1(1, 36) = 2.774$, $p=0.104$; $F_2(1, 97) = 2.549$, $p=0.114$]. This tendency was stronger in nonword targets by participants [$F_1(1, 36) = 3.681$, $p=0.063$] and significant by items [$F_2(1, 98) = 4.877$, $p=0.030$, $f=0.22$]. The interaction is due to a significantly larger congruency effect with used primes compared to novel primes.

Discussion

The stimulus-response mapping account makes a very strong prediction by attributing the effect to used primes only. In contrast, the data showed that the priming effect was present in novel primes as well. The difference in the elicited congruency effect between novel and used primes were very small with about 5 ms. In word targets the effect was numerically even larger with novel primes than with used primes. If at all, this is in contrast to the hypothesis that congruency effects in Experiment 2 were the result of stimulus-response mappings. But there was indication that the congruency effect in nonword targets was larger with used than with novel primes. Importantly, the data showed that the congruency effect was present in nonword targets with novel primes, where the effect size was almost identical across novel and used primes. Thus, the obtained congruency priming effect could not be attributed to stimulus-response mappings.

A.2 Adaptation

The analysis of prime novelty showed that there was a tendency for congruency and prime novelty to interact. In word targets this refers to a stronger congruency effect with novel primes compared to used primes, whereas in nonword targets a stronger congruency effect occurred in used primes compared to novel primes. An alternative explanation for the numerical effects in RT and the tendency in error scores was that participants adapted to the experimental task during the session. This assumption was supported by a main effect of prime novelty indicating faster responses to trials with used primes than to trials with novel primes. In the course of the experiment novel

primes were more likely at the beginning where responses were expected to be slower, whereas used primes were more likely towards the end of the session and responses were faster. A similar explanation holds with regards to error scores. If participants tuned into the task, they are expected to become more sensitive to the primes (or the type of primes) towards the end of the session where most primes were used. Thus, an interaction as reported above could emerge even though there is no contribution of stimulus-response mappings. A post-hoc analysis was performed to test the hypothesis of adaptation.

Analysis

In a post-hoc analysis all data points were tagged according to the block of the experiment in which they appeared. The results are shown in Table A.3 for word and Table A.4 for nonword targets. Thus, the values for the newly created block variable ranged from *1* to *4*. A more fine grained analysis had to be excluded, because the matrix of data points would have become very sparse. Hence, the following analysis included congruency and block as a repeated factor and list as a random factor. If this analysis shows that participants became faster during the experiment, it strongly suggests that the reason for finding a main effect of prime novelty is due to a speeding up process during the experiment.

Results

The results of the analysis showed again a strong main effect of congruency in word and nonword targets being in line with all prior analyses. The main effect of block, indicating a speeding up during the experiment, was not significant in word targets in the analysis by participants [$F_1(3, 108) = 2.470$, $p=0.062$], but it was by items [$F_2(3, 276) = 4.815$, $p=0.003$, $f=0.23$]. A post-hoc test showed that participants responded significantly slower in first block than in the second and fourth (final) block (all *p*s<0.050). There was a tendency for this difference between the first and the third block ($p<0.150$). In nonword targets the main effect of block was significant [$F_1(1, 36) = 6.215$, $p=0.001$, $f=0.42$; $F_2(3, 285) = 8.212$, $p<0.001$, $f=0.29$]. A post-hoc test showed the same result as found in word targets. The responses in the first block were significantly slower than the second and fourth block in all analyses (all *p*s<0.050).

There were tendencies for the difference between first and the third block (*p*s<0.200). There was no interaction of congruency and block in word targets [$F_1(3, 108) = 0.285$, $p=0.836$; $F_2(3, 276) = 1.127$, $p=0.339$]. This interaction was not significant in nonword targets as well [$F_1(3, 108) = 0.930$, $p=0.429$; $F_2(3, 276) = 1.466$, $p=0.224$].

The analysis of error scores showed a main effect of congruency and was in line with previous analyses. There was no main effect of block in word [$F_1(3, 108) = 0.801$, $p=0.496$; $F_2(3, 279) = 1.688$, $p=0.170$] and nonword targets [$F_1(3, 108) = 1.730$, $p=0.165$; $F_2(3, 285) = 1.829$, $p=0.142$]. Furthermore, there was no interaction between congruency and block in words [$F_1(3, 108) = 1.340$, $p=0.265$; $F_2(3, 279) = 1.191$, $p=0.313$] or in nonwords [$F_1(3, 108) = 1.778$, $p=0.156$; $F_2(3, 285) = 0.425$, $p=0.735$].

Discussion

The analysis showed that the first block was the slowest in the experimental session. Participants speeded up during the first block and there were only small differences between the average speed in the following blocks. This can be attributed to tuning into the task, e.g. adapting the thresholds or shifting weight to most reliable source of

Table A.3: Mean reaction times in ms and error ratio of Experiment 2 in Word targets in percent as a function of congruency and block

	Block							
	1		2		3		4	
Congruency	RT	Error	RT	Error	RT	Error	RT	Error
Incongruent	510	3.00	493	3.08	498	4.25	492	4.95
Congruent	481	1.81	472	2.39	473	2.18	476	2.15
Effect	29	1.19	21	0.69	25	2.07	16	2.80

Table A.4:Mean reaction times in ms and error ratio of Experiment 2 in Nonword targets in percent as a function of congruency and block

	Block							
	1		2		3		4	
Congruency	RT	Error	RT	Error	RT	Error	RT	Error
Incongruent	533	2.19	514	3.71	524	4.45	509	4.78
Congruent	513	1.07	498	3.01	493	2.13	486	1.41
Effect	20	1.12	16	0.70	31	2.32	23	3.37

information. As a result of the experimental procedure the first block comprised the largest number of novel primes. Thus, the main effect of prime novelty can also be attributed to participants speeding up and adapting during the experimental session. The absence of an interaction between congruency and block indicates that the congruency effect was not affected by general speeding up of the participants.

A.3 General Discussion

The post-hoc analyses of prime novelty and block showed that stimulus-response mapping is unlikely to be the mechanism underlying the congruency effects reported in Experiment 2 and that participants adapted during the experimental session, especially during the first block. In this experimental procedure the novel primes occurred more frequent early in the experiment than towards the end of the session. Thus, an effect of prime novelty could emerge even though it actually reflects an effect of adaptation. There were some numerical differences of the congruency effect to interact with prime novelty, but these differences were small. Most importantly, there was a congruency effect with novel primes which was not expected in a stimulus-response mapping account. The effect size was larger with novel primes in word targets and smaller in nonword targets. Hence, stimulus-response mapping was unlikely in Experiment 2. This finding was compatible with other studies that reported effects with novel primes (e.g., Greenwald et al., 2003; Kinoshita & Hunt, 2008; Kinoshita & Norris, 2010; Naccache & Dehaene, 2001; Perea & Gotor, 1997; Reynvoet et al., 2005).

Appendix B: Experimental stimuli

The stimuli of the unprimed lexical decision experiment (Experiment 1) are listed by their lexicality, pronounceability and typicality. All stimuli that were used in the masked primed lexical decision experiments that were testing response congruency (Experiment 2 to 8) are presented as a triplet of the form *<TARGET, congruent prime, incongruent prime>*. In Experiment 9 and 10 the triplets are ordered as *<TARGET, high typicality prime, low typicality prime>*.

B.1 Stimuli of Experiment 1

Word targets

High typicality: cover; theme; quest; might; count; trout; cough; print; bound; yeast; spent; shout; ought; stand; eight; front; fight; right; tract; about; thigh; light; their; could; sight; there; agent; crest; force; sport; saint; after; court; thing; offer; prove; press; wound; event; would; expert; shaver; action; threat; expand; buying; should; recent; height; common; though; plight; budget; extent; proper; report; decent; commit; strand; format; priest; latent; office; lament; caught; market; bright; export; notion; expect; weight; fright; invest; prompt; forest; forget; accent; reject; intent; invent

Low typicality: vinyl; soggy; relax; grasp; epoch; flaky; noisy; spiky; sneak; enemy; telex; dodgy; idiom; aloud; tawny; idyll; codex; twirl; snipe; gauze; check; elbow; index; swirl; tulip; clerk; ocean; enjoy; error; smirk; kiosk; zippy; slimy; daisy; xerox; swamp; zebra; yummy; odour; topaz; baffle; eyelid; ethnic; hungry; icebox; chunky; snooze; hyphen; groovy; pigeon; cowboy; melody; zigzag; fluffy; unripe; stupid; zombie; jigsaw; occupy; reflex; loudly; kidnap; filthy; typify; galaxy; wisdom; shrunk; sleepy; nephew; nutmeg; pianos; piglet; clumsy; pyjama; matrix; cortex; rhythm; embryo; asthma; cognac

Pronounceable nonwords

High typicality: reent; thaid; unces; inges; conat; derat; recto; inced; eques; thels; monat; inals; ficat; perat; accut; yeent; siont; iment; thets; istat; proust; foreas; frount; stiont; sument; uncent; whount; wition; theigh; itiont; offies; ationt; prould; frould; weents; prount; itions; govent; thount; thaves

Low typicality: imgoz; syheu; hytyk; zivix; zobix; zhofo; kobgu; xarvy; ubopa; atjox; ojuez; zuavu; udcun; snoaz; uvoap; udnop; ewusy; ofoiz; owofl; aykuz; eznotz; umdaxe; xibtib; vobduy; namgox; oyfoby; zoruyu; sloxuk; zezlol; tusfoh; omtelz; kymtap; paipyz; igvoik; ghikik; ujumli; vimfio; ikupuz; robkox; kufylt

Unpronounceable nonwords

High typicality: acque; shmat; czent; offsh; ithre; doeir; stmes; afthe; thdat; excht; chnit; thlys; mrsto; dexch; exchn; stmet; afght; thfut; chnis; intst; straqi; thment; chrent; ingthe; rafght; efught; exchis; minght; thadve; eaccut; forpre; racces; aciand; thwass; safght; yughts; thdrat; thques; whighe; stment

Low typicality: baetg; hrspo; rgupg; crvoz; rpapm; fdusq; vlpej; nkwog; itrnl; mgoav; agtmb; rpysi; pciqu; lneej; dfowo; tcytl; tcumw; iedbh; mntuc; ljemw; meoatj; gcolqa; xegoue; wlajma; zayfop; eauawr; aoykrn; ndskyj; smkhml; oxdojl; mpkosq; iesdss; noeogu; megshp; teaoac; ahptuh; hrhtel; ynciuf; eciiuk; clcanl

B.2 Stimuli of Experiment 2

Word targets

ABOUT, price, phgil; ABOVE, child, dcpsh; ADMIT, crown, lprph; AGAIN, press, wzzyc; ALLOW, trust, empxn; BEACH, point, sugfx; BEGIN, start, suvsc; BEING, local, rrzsh; BELOW, thing, xvump; BRAIN, judge, uckos; BREAD, count, migxu; BRING, plate, cljos; BROAD, night, sugln; BROWN, cause, fgipe; BUILD, front, acgph; CATCH, order, voysl; CAUSE, limit, migpv; CHILD, often, auqjs; CLASS, worth, djekb; CLEAR, visit, miguz; CLOSE, grand, avkra; COACH, under, fdpei; COAST, bring, dqrki; COULD, right, jazjh;

COUNT, large, sykdd; COURT, spend, syilb; COVER, small, pujuw; CROWN, quiet, miytd; EARLY, sight, cigii; EQUAL, short, ziwzw; EVENT, class, ocsym; EVERY, match, chnzk; FIGHT, power, embxz; FIRST, leave, hkaub; FORCE, admit, apvxa; FRONT, equal, celqi; FULLY, being, wzxoo; GRAND, whole, mjymb; GREEN, would, abwwa; HENCE, party, aivpg; HOUSE, plain, jaxxa; JUDGE, still, syniy; LARGE, youth, ftnzi; LAUGH, speed, mzdpe; LEAVE, sound, roccp; LIGHT, broad, eacyd; LIMIT, every, dazko; LOCAL, three, suuhp; MAJOR, hence, fegyp; MARCH, event, nyxoo; MATCH, world, spnfr; MIGHT, cover, safcn; NEVER, coast, gihlw; NIGHT, close, rojul; OFFER, plant, bxyia; OFTEN, build, mihjw; ORDER, laugh, sujcw; OTHER, again, ncllb; PARTY, house, ievzb; PLACE, union, mnstw; PLAIN, other, bfmsh; PLANT, where, ricjg; PLATE, young, igicw; POINT, march, hfdef; POUND, clear, qilil; POWER, light, clyym; PRESS, might, giagj; PRICE, allow, mvsym; QUIET, brown, aklph; RIGHT, below, axayw; ROUGH, table, bvenn; ROUND, state, ecajw; SEVEN, court, axypg; SHORT, fully, audbc; SIGHT, bread, mqava; SINCE, about, agtfw; SMALL, offer, pekiq; SOUND, catch, micpj; SPEED, brain, acnhc; SPEND, major, chriv; STAND, force, miqxp; START, pound, cungd; STATE, rough, imkzo; STILL, coach, hcneo; TABLE, round, rojrp; THICK, early, saqzs; THING, above, rakpw; THREE, could, cxnio; TRUST, beach, hymbp; UNDER, thick, oyizi; UNION, watch, sxzel; VISIT, green, cgkra; WATCH, never, miqyd; WHERE, stand, oljul; WHOLE, first, fbdfi; WORLD, begin, imimj; WORTH, seven, cgglb; WOULD, fight, icsxa; YOUNG, place, raxri; YOUTH, since, dvnel

Nonword targets

AXAYW, ievzb, build; SUJCW, imkzo, right; APVXA, ricjg, house; ZIWZW, apvxa, bread; MIQYD, sugln, other; SUGFX, acnhc, thick; MVSYM, lprph, laugh; AGTFW, hcneo, could; RAKPW, suvsc, light; AXYPG, ftnzi, would; CGKRA, oljul, seven; ECAJW, mvsym, sight; SAFCN, miqyd, below; OLJUL, bvenn, since; SAQZS, hymbp, bring; CHNZK, giagj, admit; MIQXP, wzzyc, where; CUNGD, bxyia, limit; SXZEL, avkra, march; CLYYM, rrzsh, three; CGGLB, mihjw, under; ROJRP, celqi, again; AUQJS, cgglb, every; IEVZB, rakpw, world; HCNEO, qilil, first; BVENN, axayw, short; SUUHP, ecajw, early; MIYTD, sujcw, coach; IMKZO, dvnel, party; PHGIL, embxz, front; FTNZI, cgkra, local; FEGYP, cxnio, brain; DQRKI, voysl, plate; AIVPG, uckos, sound; FDPEI, chnzk, about; JAXXA, gihlw, union; NYXOO, dcpsh, class; XVUMP, cljos, clear; MIGPV, hkaub, beach; SYKDD, empxn, match;

SUVSC, fgipe, power; DVNEL, micpj, catch; ABWWA, roccp, young; ACGPH, sykdd, offer; MQAVA, pujuw, worth; CIGII, mjymb, major; SUGLN, jazjh, broad; FBDFI, rojrp, press; PEKIQ, clyym, round; RRZSH, agtfw, above; MIGXU, aklph, close; RAXRI, cungd, pound; DCPSH, abwwa, often; IMIMJ, ncllb, count; HKAUB, ocsym, visit; QILIL, nyxoo, force; EMBXZ, cigii, start; DAZKO, fegyp, fully; WZZYC, fbdfi, plant; MICPJ, hfdef, green; GIHLW, dazko, court; SYILB, wzxoo, event; MZDPE, sugfx, allow; HYMBP, auqjs, cause; MIHJW, eacyd, grand; RICJG, mzdpe, youth; FGIPE, mqava, watch; OYIZI, djekb, place; ACNHC, dqrki, judge; AKLPH, miytd, order; JAZJH, syilb, point; EACYD, miguz, rough; ICSXA, rojul, never; CXNIO, bfmsh, table; GIAGJ, xvump, brown; MIGUZ, safcn, spend; LPRPH, icsxa, coast; VOYSL, migxu, hence; WZXOO, syniy, might; UCKOS, ziwzw, plain; CHRIV, axypg, state; BXYIA, spnfr, whole; ROCCP, jaxxa, begin; AUDBC, miqxp, still; PUJUW, chriv, being; MNSTW, acgph, price; BFMSH, oyizi, equal; ROJUL, migpv, thing; OCSYM, aivpg, fight; EMPXN, saqzs, child; HFDEF, raxri, crown; AVKRA, imimj, night; DJEKB, suuhp, small; CLJOS, pekiq, quiet; IGICW, sxzel, speed; NCLLB, fdpei, trust; SYNIY, audbc, cover; CELQI, mnstw, stand; MJYMB, phgil, large; SPNFR, igicw, leave

B.3 Stimuli of Experiment 3

Word targets

ABOUT, since, phgil; ABOVE, thing, dcpsh; ADMIT, force, lprph; AGAIN, other, wzzyc; ALLOW, price, empxn; BEACH, trust, sugfx; BEGIN, world, suvsc; BEING, fully, rrzsh; BELOW, right, xvump; BRAIN, speed, uckos; BREAD, sight, migxu; BRING, coast, cljos; BROAD, light, sugln; BROWN, quiet, fgipe; BUILD, often, acgph; CATCH, sound, voysl; CAUSE, brown, migpv; CHILD, above, auqjs; CLASS, event, djekb; CLEAR, pound, miguz; CLOSE, night, avkra; COACH, still, fdpei; COAST, never, dqrki; COULD, three, jazjh; COUNT, bread, sykdd; COURT, seven, syilb; COVER, might, pujuw; CROWN, admit, miytd; EARLY, thick, cigii; EQUAL, front, ziwzw; EVENT, march, ocsym; EVERY, limit, chnzk; FIGHT, would, embxz; FIRST, whole, hkaub; FORCE, stand, apvxa; FRONT, build, celqi; FULLY, short, wzxoo; GRAND, close, mjymb; GREEN, visit, abwwa; HENCE, major, aivpg; HOUSE, party, jaxxa; JUDGE, brain, syniy; LARGE, count, ftnzi; LAUGH, order, mzdpe; LEAVE, first, roccp; LIGHT, power, eacyd; LIMIT, cause, dazko; LOCAL, being, suuhp; MAJOR, spend, fegyp; MARCH, point, nyxoo; MATCH, every, spnfr; MIGHT,

press, safcn; NEVER, watch, gihlw; NIGHT, broad, rojul; OFFER, small, bxyia; OFTEN, child, mihjw; ORDER, catch, sujcw; OTHER, plain, ncllb; PARTY, hence, ievzb; PLACE, young, mnstw; PLAIN, house, bfmsh; PLANT, offer, ricjg; PLATE, bring, igicw; POINT, beach, hfdef; POUND, start, qilil; POWER, fight, clyym; PRESS, again, giagj; PRICE, about, mvsym; QUIET, crown, aklph; RIGHT, could, axayw; ROUGH, state, bvenn; ROUND, table, ecajw; SEVEN, worth, axypg; SHORT, equal, audbc; SIGHT, early, mqava; SINCE, youth, agtfw; SMALL, cover, pekiq; SOUND, leave, micpj; SPEED, laugh, acnhc; SPEND, court, chriv; STAND, where, miqxp; START, begin, cungd; STATE, round, imkzo; STILL, judge, hcneo; TABLE, rough, rojrp; THICK, under, saqzs; THING, below, rakpw; THREE, local, cxnio; TRUST, allow, hymbp; UNDER, coach, oyizi; UNION, place, sxzel; VISIT, clear, cgkra; WATCH, union, miqyd; WHERE, plant, oljul; WHOLE, grand, fbdfi; WORLD, match, imimj; WORTH, class, cgglb; WOULD, green, icsxa; YOUNG, plate, raxri; YOUTH, large, dvnel

Nonword targets

ABLIL, qyvsj, court; ABLUN, vfrsz, offer; AFFSA, jhwpz, power; APLIA, jrvgq, other; ARLUD, jcxsz, being; AWYNS, qcglj, could; BAKKA, xhmxv, price; BIFEL, xpkhk, march; BILIR, kveuq, coast; BLEPH, zufdk, trust; BLICO, qfyqz, speed; BULUM, qqcwz, catch; CEENA, xgpvq, would; COCRI, qqgsq, leave; COMLI, zghgz, seven; CULGA, xnrxq, first; DEDIO, xjtjv, plant; DULEW, ztcik, start; DULUS, xrjpj, major; ECHUR, qivkj, limit; EFEBA, zpnrv, thing; EGESK, xuqvq, union; EGRIS, qdbnq, youth; ENKIS, xrrfv, rough; ENRYS, kjkdx, coach; EUMPE, xakwx, watch; FEIRA, qzdtq, young; FESPE, qqoqq, broad; FICIM, xsptj, early; FILDI, qbwkk, house; FLEUM, kkjhz, party; FLIKI, vmxjx, press; FOCRO, zswvz, night; FOICT, xbpdx, judge; FULPH, qtjxq, since; FUTEW, qqblq, again; GEDUA, zhnhj, still; GEDUM, jqpsz, crown; GIBIR, vjtdz, place; GIBUR, zvpfx, whole; HUDEP, jyxcq, local; HYRIS, xcpxk, table; IGISS, qvvpx, order; ILDIC, ktkrz, about; ILVIS, zhcjz, bread; IVIVE, kqpwx, short; JACUM, qigtk, spend; JAZDA, qfpyk, quiet; JOVIC, xstmz, where; JUDYA, zfkgx, cover; JULIM, xctbz, never; KAKEM, xydjx, count; KARJU, vwffx, point; KAURS, xmqdx, below; KEDDS, vpjjx, match; KESUR, xxvyj, fight; KLUDS, jwnrx, often; KNUCH, jdjwq, plate; KOCIP, zhnuz, equal; KULDE, xvgwz, front; LEFEB, qhnrj, sound; LEFOL, zcrkv, stand; LIKIR, qzvyj, cause; LIZIM, xdcbx, under; LUSEL, jykcz, brown; MIKAK, zsqvz, world; MOKIL, xfpvv, state; NAULF, qhdjk, visit; NEFEB, zulmk,

class; NEZAL, xyjxx, worth; OBSEP, zukuj, laugh; OCOMP, kwsuz, large; OUCIS, kdybq, every; OUNEM, xrrlv, light; PESPE, xdqwx, grand; PESUR, zwdtj, might; PLEVO, kbtdq, sight; POCEF, xlmhx, right; ROSCO, zdeuj, admit; ROSUR, zgwfk, child; RUILM, xcxpk, hence; RUKIL, jybjq, beach; SARRO, zfekv, begin; SAULP, xcbkk, force; SEPRI, qbnnj, allow; SIBOU, zdpkj, clear; SIFLU, vjggz, three; SKAKA, qgmhq, round; SUROK, qzcez, plain; SURSA, zwmgk, event; SYSLI, qfqhx, above; TEIRP, xncnq, fully; TESEM, zbbrq, build; TUECT, xkyik, brain; UNENG, xyrfj, thick; URVIR, zhbfj, small; VARLI, jsydz, pound; VUKET, zwlwj, bring; WABIA, xkxrx, close; ZOFLI, xsnxq, green

B.4 Stimuli of Experiment 4

Word targets

ALERT, sound, cungd; ATLAS, offer, empxn; AVAIL, green, sugfx; BADGE, visit, pujuw; BINGO, place, jaxxa; BISON, march, wzzyc; BLACK, other, suuhp; BLITZ, young, hcneo; BLUFF, three, axypg; BLUNT, force, miqxp; BOARD, hence, sugln; BOOZE, plain, suvsc; BROWN, thick, sxzel; BULGE, coast, mihjw; CARGO, spend, miqyd; CHARM, often, oyizi; CLASS, judge, rojrp; CLOUD, brain, giagj; COACH, fully, pekiq; COBRA, thing, fgipe; COCOA, being, miytd; CROOK, plant, fbdfi; CROWD, again, fegyp; CRUDE, allow, migpv; CRUMB, whole, syniy; CYCLE, round, avkra; DELTA, rough, cigii; DWARF, close, cxnio; EAGLE, count, dqrki; ERASE, fight, imkzo; EXTRA, could, dcpsh; FAULT, price, ocsym; FIELD, major, uckos; FLICK, house, apvxa; FLUTE, broad, rrzsh; FRAUD, below, imimj; GAUGE, world, lprph; GROUP, beach, syilb; HAZEL, brown, roccp; HEDGE, about, oljul; ICING, party, djekb; INERT, local, clyym; KNIFE, coach, xvump; KNOCK, speed, bfmsh; LEVEL, admit, axayw; LOCAL, press, ievzb; LODGE, start, icsxa; LOOSE, grand, abwwa; LOTUS, begin, acgph; LUNCH, power, agtfw; LYRIC, state, wzxoo; MIMIC, plate, aklph; MINOR, table, hkaub; MORAL, sight, ziwzw; MOTOR, laugh, ecajw; MOTTO, bring, safcn; MOURN, leave, fdpei; NURSE, child, gihlw; OBESE, match, micpj; OZONE, first, phgil; PANDA, where, celqi; PANEL, short, sykdd; PANIC, every, voysl; PIECE, worth, rojul; PILOT, cause, auqjs; PIZZA, court, mnstw; PLAZA, front, sujcw; PROUD, might, ncllb; QUAKE, right, chriv; RAPID, seven, cljos; REIGN, watch, dazko; RIDGE, small, acnhc; RIFLE, youth, nyxoo; ROUGH, class, mqava; SCOOP, bread, ftnzi; SCRAP, night, mjymb; SCRUB, event, aivpg; SOLID, never, hymbp; SPACE, limit, migxu; SPASM, quiet,

chnzk; SPOON, large, jazjh; STAFF, build, mzdpe; STUNT, cover, raxri; STYLE, union, ricjg; SWARM, point, dvnel; SWIRL, above, audbc; SYRUP, catch, embxz; THEFT, would, mvsym; THORN, equal, saqzs; THUMB, order, qilil; TIMID, clear, rakpw; TOTAL, under, bvenn; TULIP, crown, eacyd; TWIST, early, hfdef; VENUE, light, cgkra; VODKA, since, cgglb; VOICE, stand, spnfr; WATER, pound, miguz; WHEEL, trust, bxyia; ZEBRA, still, igicw

Nonword targets

ABWWA, roccp, young; ACGPH, sykdd, offer; ACNHC, dqrki, judge; AGTFW, hcneo, could; AIVPG, uckos, sound; AKLPH, miytd, order; APVXA, ricjg, house; AUDBC, miqxp, still; AUQJS, cgglb, every; AVKRA, imimj, night; AXAYW, ievzb, build; AXYPG, ftnzi, would; BFMSH, oyizi, equal; BVENN, axayw, short; BXYIA, spnfr, whole; CELQI, mnstw, stand; CGGLB, mihjw, under; CGKRA, oljul, seven; CHNZK, giagj, admit; CHRIV, axypg, state; CIGII, mjymb, major; CLJOS, pekiq, quiet; CLYYM, rrzsh, three; CUNGD, bxyia, limit; CXNIO, bfmsh, table; DAZKO, fegyp, fully; DCPSH, abwwa, often; DJEKB, suuhp, small; DQRKI, voysl, plate; DVNEL, micpj, catch; EACYD, miguz, rough; ECAJW, mvsym, sight; EMBXZ, cigii, start; EMPXN, saqzs, child; FBDFI, rojrp, press; FDPEI, chnzk, about; FEGYP, cxnio, brain; FGIPE, mqava, watch; FTNZI, cgkra, local; GIAGJ, xvump, brown; GIHLW, dazko, court; HCNEO, qilil, first; HFDEF, raxri, crown; HKAUB, ocsym, visit; HYMBP, auqjs, cause; ICSXA, rojul, never; IEVZB, rakpw, world; IGICW, sxzel, speed; IMIMJ, ncllb, count; IMKZO, dvnel, party; JAXXA, gihlw, union; JAZJH, syilb, point; LPRPH, icsxa, coast; MICPJ, hfdef, green; MIGPV, hkaub, beach; MIGUZ, safcn, spend; MIGXU, aklph, close; MIHJW, eacyd, grand; MIQXP, wzzyc, where; MIQYD, sugln, other; MIYTD, sujcw, coach; MJYMB, phgil, large; MNSTW, acgph, price; MQAVA, pujuw, worth; MVSYM, lprph, laugh; MZDPE, sugfx, allow; NCLLB, fdpei, trust; NYXOO, dcpsh, class; OCSYM, aivpg, fight; OLJUL, bvenn, since; OYIZI, djekb, place; PEKIQ, clyym, round; PHGIL, embxz, front; PUJUW, chriv, being; QILIL, nyxoo, force; RAKPW, suvsc, light; RAXRI, cungd, pound; RICJG, mzdpe, youth; ROCCP, jaxxa, begin; ROJRP, celqi, again; ROJUL, migpv, thing; RRZSH, agtfw, above; SAFCN, miqyd, below; SAQZS, hymbp, bring; SPNFR, igicw, leave; SUGFX, acnhc, thick; SUGLN, jazjh, broad; SUJCW, imkzo, right; SUUHP, ecajw, early; SUVSC, fgipe, power; SXZEL, avkra, march; SYILB, wzxoo, event; SYKDD, empxn, match; SYNIY, audbc, cover; UCKOS, ziwzw, plain; VOYSL, migxu,

hence; WZXOO, syniy, might; WZZYC, fbdfi, plant; XVUMP, cljos, clear; ZIWZW, apvxa, bread

B.5 Stimuli of Experiment 5

Word targets

ALERT, sound, cungd; ATLAS, offer, empxn; AVAIL, green, sugfx; BADGE, visit, pujuw; BINGO, place, jaxxa; BISON, march, wzzyc; BLACK, other, suuhp; BLITZ, young, hcneo; BLUFF, three, axypg; BLUNT, force, miqxp; BOARD, hence, sugln; BOOZE, plain, suvsc; BROWN, thick, sxzel; BULGE, coast, mihjw; CARGO, spend, miqyd; CHARM, often, oyizi; CLASS, judge, rojrp; CLOUD, brain, giagj; COACH, fully, pekiq; COBRA, thing, fgipe; COCOA, being, miytd; CROOK, plant, fbdfi; CROWD, again, fegyp; CRUDE, allow, migpv; CRUMB, whole, syniy; CYCLE, round, avkra; DELTA, rough, cigii; DWARF, close, cxnio; EAGLE, count, dqrki; ERASE, fight, imkzo; EXTRA, could, dcpsh; FAULT, price, ocsym; FIELD, major, uckos; FLICK, house, apvxa; FLUTE, broad, rrzsh; FRAUD, below, imimj; GAUGE, world, lprph; GROUP, beach, syilb; HAZEL, brown, roccp; HEDGE, about, oljul; ICING, party, djekb; INERT, local, clyym; KNIFE, coach, xvump; KNOCK, speed, bfmsh; LEVEL, admit, axayw; LOCAL, press, ievzb; LODGE, start, icsxa; LOOSE, grand, abwwa; LOTUS, begin, acgph; LUNCH, power, agtfw; LYRIC, state, wzxoo; MIMIC, plate, aklph; MINOR, table, hkaub; MORAL, sight, ziwzw; MOTOR, laugh, ecajw; MOTTO, bring, safcn; MOURN, leave, fdpei; NURSE, child, gihlw; OBESE, match, micpj; OZONE, first, phgil; PANDA, where, celqi; PANEL, short, sykdd; PANIC, every, voysl; PIECE, worth, rojul; PILOT, cause, auqjs; PIZZA, court, mnstw; PLAZA, front, sujcw; PROUD, might, ncllb; QUAKE, right, chriv; RAPID, seven, cljos; REIGN, watch, dazko; RIDGE, small, acnhc; RIFLE, youth, nyxoo; ROUGH, class, mqava; SCOOP, bread, ftnzi; SCRAP, night, mjymb; SCRUB, event, aivpg; SOLID, never, hymbp; SPACE, limit, migxu; SPASM, quiet, chnzk; SPOON, large, jazjh; STAFF, build, mzdpe; STUNT, cover, raxri; STYLE, union, ricjg; SWARM, point, dvnel; SWIRL, above, audbc; SYRUP, catch, embxz; THEFT, would, mvsym; THORN, equal, saqzs; THUMB, order, qilil; TIMID, clear, rakpw; TOTAL, under, bvenn; TULIP, crown, eacyd; TWIST, early, hfdef; VENUE, light, cgkra; VODKA, since, cgglb; VOICE, stand, spnfr; WATER, pound, miguz; WHEEL, trust, bxyia; ZEBRA, still, igicw

Nonword targets

ABLIL, roccp, court; ABLUN, migpv, offer; AFFSA, pekiq, power; APLIA, mnstw, other; ARLUD, cigii, being; AWYNS, cgglb, could; BAKKA, nyxoo, price; BIFEL, jazjh, march; BILIR, fegyp, coast; BLEPH, sykdd, trust; BLICO, suuhp, speed; BULUM, chriv, catch; CEENA, oljul, would; COCRI, sugfx, leave; COMLI, hfdef, seven; CULGA, miqyd, first; DEDIO, xvump, plant; DULEW, aivpg, start; DULUS, hcneo, major; ECHUR, syniy, limit; EFEBA, imkzo, thing; EGESK, cxnio, union; EGRIS, ncllb, youth; ENKIS, clyym, rough; ENRYS, giagj, coach; EUMPE, dqrki, watch; FEIRA, sugln, young; FESPE, dazko, broad; FICIM, rojul, early; FILDI, suvsc, house; FLEUM, avkra, party; FLIKI, mvsym, press; FOCRO, imimj, night; FOICT, auqjs, judge; FULPH, miytd, since; FUTEW, icsxa, again; GEDUA, ocsym, still; GEDUM, cljos, crown; GIBIR, empxn, place; GIBUR, eacyd, whole; HUDEP, syilb, local; HYRIS, agtfw, table; IGISS, aklph, order; ILDIC, rojrp, about; ILVIS, mzdpe, bread; IVIVE, uckos, short; JACUM, fgipe, spend; JAZDA, lprph, quiet; JOVIC, rakpw, where; JUDYA, embxz, cover; JULIM, safcn, never; KAKEM, ricjg, count; KARJU, oyizi, point; KAURS, celqi, below; KEDDS, bxyia, match; KESUR, hymbp, fight; KLUDS, igicw, often; KNUCH, sxzel, plate; KOCIP, bvenn, equal; KULDE, axayw, front; LEFEB, migxu, sound; LEFOL, chnzk, stand; LIKIR, sujcw, cause; LIZIM, pujuw, under; LUSEL, mjymb, brown; MIKAK, voysl, world; MOKIL, spnfr, state; NAULF, ievzb, visit; NEFEB, axypg, class; NEZAL, miqxp, worth; OBSEP, mqava, laugh; OCOMP, ftnzi, large; OUCIS, jaxxa, every; OUNEM, abwwa, light; PESPE, hkaub, grand; PESUR, fbdfi, might; PLEVO, raxri, sight; POCEF, qilil, right; ROSCO, ziwzw, admit; ROSUR, ecajw, child; RUILM, acnhc, hence; RUKIL, dcpsh, beach; SARRO, micpj, begin; SAULP, djekb, force; SEPRI, audbc, allow; SIBOU, wzzyc, clear; SIFLU, acgph, three; SKAKA, miguz, round; SUROK, fdpei, plain; SURSA, mihjw, event; SYSLI, cgkra, above; TEIRP, bfmsh, fully; TESEM, gihlw, build; TUECT, phgil, brain; UNENG, rrzsh, thick; URVIR, wzxoo, small; VARLI, cungd, pound; VUKET, saqzs, bring; WABIA, dvnel, close; ZOFLI, apvxa, green

B.6 Stimuli of Experiment 6

Word targets

ALERT, mimic, oucis; ATLAS, ozone, echur; AVAIL, nurse, gedum; BADGE, flick, ocomp; BINGO, scrap, sursa; BISON, theft, vuket; BLACK, inert, ounem; BLITZ, swarm, ougam; BLUFF, panda, rosco; BLUNT, voice, hyris; BOARD, cycle, ilvis; BOOZE, stunt, karju; BROWN, field, fulph; BULGE, rapid, jovic; CARGO, level, kulde; CHARM, booze, lusel; CLASS, thorn, dedio; CLOUD, staff, egesk; COACH, reign, pesur; COBRA, style, kedds; COCOA, thumb, enrys; CROOK, blunt, uneng; CROWD, gauge, ablun; CRUDE, blitz, zilys; CRUMB, panel, affsa; CYCLE, rough, sibou; DELTA, bingo, gibur; DWARF, obese, eumpe; EAGLE, mourn, kocip; ERASE, pilot, fildi; EXTRA, scoop, mokil; FAULT, minor, egris; FIELD, group, sarro; FLICK, erase, gedua; FLUTE, board, jazda; FRAUD, loose, sysli; GAUGE, swirl, likir; GROUP, class, feiza; HAZEL, proud, urvir; HEDGE, cobra, varli; ICING, alert, plevo; INERT, plaza, bakka; KNIFE, cocoa, rosur; KNOCK, badge, lefeb; LEVEL, knock, surok; LOCAL, hedge, teirp; LODGE, charm, wabia; LOOSE, dwarf, judya; LOTUS, ridge, efeba; LUNCH, pizza, ivive; LYRIC, venue, tesem; MIMIC, extra, obsep; MINOR, black, pespe; MORAL, icing, igiss; MOTOR, wheel, nezal; MOTTO, scrub, culga; MOURN, space, aplia; NURSE, timid, mikak; OBESE, fault, rukil; OZONE, lyric, bilir; PANDA, flute, ficim; PANEL, twist, cocri; PANIC, lodge, kesur; PIECE, brown, dulus; PILOT, zebra, awyns; PIZZA, lotus, lefol; PLAZA, bison, enkis; PROUD, eagle, ceena; QUAKE, motor, lizim; RAPID, motto, tuect; REIGN, local, couza; RIDGE, lunch, bulum; RIFLE, coach, jacum; ROUGH, panic, fliki; SCOOP, rifle, arlud; SCRAP, bulge, bifel; SCRUB, delta, zofli; SOLID, crumb, kakem; SPACE, bluff, julim; SPASM, crook, futew; SPOON, water, ablil; STAFF, crude, dulew; STUNT, cargo, gibir; STYLE, fraud, burac; SWARM, piece, hudep; SWIRL, quake, nefeb; SYRUP, total, blico; THEFT, moral, ildic; THORN, avail, fespe; THUMB, vodka, sepri; TIMID, hazel, kaurs; TOTAL, knife, knuch; TULIP, crowd, focro; TWIST, cloud, bleph; VENUE, spasm, comli; VODKA, syrup, fleum; VOICE, atlas, saulp; WATER, solid, kluds; WHEEL, spoon, skaka; ZEBRA, tulip, foict

Nonword targets

ABLIL, gedum, ozone; ABLUN, jovic, ridge; AFFSA, kulde, lunch; APLIA, kedds, crude; ARLUD, egesk, voice; AWYNS, bulum, crumb; BAKKA, dulus, rough; BIFEL, ougam,

crowd; BILIR, uneng, cocoa; BLEPH, kaurs, mourn; BLICO, hudep, extra; BULUM, sepri, vodka; BURAC, dedio, loose; CEENA, gibur, tulip; COCRI, nefeb, theft; COMLI, pesur, quake; COUZA, egris, swirl; CULGA, enkis, timid; DEDIO, saulp, black; DULEW, ocomp, mimic; DULUS, kakem, reign; ECHUR, lizim, staff; EFEBA, igiss, solid; EGESK, varli, fault; EGRIS, fulph, coach; ENKIS, arlud, cargo; ENRYS, culga, plaza; EUMPE, affsa, class; FEIZA, surok, proud; FESPE, mikak, total; FICIM, sursa, delta; FILDI, couza, syrup; FLEUM, sarro, bison; FLIKI, burac, erase; FOCRO, eumpe, badge; FOICT, karju, spasm; FULPH, cocri, water; FUTEW, blico, moral; GEDUA, mokil, lyric; GEDUM, kocip, knock; GIBIR, pespe, panel; GIBUR, lefol, spoon; HUDEP, sysli, motor; HYRIS, futew, bulge; IGISS, vuket, thorn; ILDIC, kesur, nurse; ILVIS, ounem, crook; IVIVE, rosco, lotus; JACUM, bleph, hedge; JAZDA, urvir, wheel; JOVIC, enrys, bluff; JUDYA, plevo, pilot; JULIM, ceena, obese; KAKEM, sibou, stunt; KARJU, ildic, twist; KAURS, lefeb, level; KEDDS, gibir, blunt; KESUR, ablil, avail; KLUDS, feiza, inert; KNUCH, teirp, alert; KOCIP, dulew, hazel; KULDE, awyns, pizza; LEFEB, knuch, swarm; LEFOL, judya, rapid; LIKIR, tesem, space; LIZIM, skaka, cobra; LUSEL, foict, panic; MIKAK, lusel, style; MOKIL, efeba, scrap; NEFEB, rukil, group; NEZAL, oucis, scrub; OBSEP, jacum, charm; OCOMP, zilys, gauge; OUCIS, bakka, zebra; OUGAM, ilvis, rifle; OUNEM, hyris, atlas; PESPE, zofli, brown; PESUR, fliki, flick; PLEVO, wabia, fraud; ROSCO, bifel, knife; ROSUR, ivive, blitz; RUKIL, obsep, panda; SARRO, julim, cycle; SAULP, ficim, minor; SEPRI, jazda, local; SIBOU, nezal, eagle; SKAKA, fleum, piece; SUROK, aplia, icing; SURSA, fildi, field; SYSLI, focro, booze; TEIRP, kluds, cloud; TESEM, ablun, dwarf; TUECT, bilir, bingo; UNENG, likir, scoop; URVIR, fespe, lodge; VARLI, tuect, thumb; VUKET, comli, board; WABIA, rosur, flute; ZILYS, gedua, motto; ZOFLI, echur, venue

B.7 Stimuli of Experiment 7

Word targets

ALERT, mimic, whoss; ATLAS, ozone, grome; AVAIL, nurse, thost; BADGE, flick, shist; BINGO, scrap, cluse; BISON, theft, thake; BLACK, inert, prous; BLITZ, swarm, roure; BLUFF, panda, chave; BLUNT, voice, forse; BOARD, cycle, chise; BOOZE, stunt, dight; BROWN, field, lighe; BULGE, rapid, spost; CARGO, level, quess; CHARM, booze, sould; CLASS, thorn, wough; CLOUD, staff, wasis; COACH, reign, rieve; COBRA, style, thele;

COCOA, thumb, itter; CROOK, blunt, fally; CROWD, gauge, quall; CRUDE, blitz, wasky; CRUMB, panel, thild; CYCLE, rough, nound; DELTA, bingo, pring; DWARF, obese, becon; EAGLE, mourn, prind; ERASE, pilot, yough; EXTRA, scoop, shing; FAULT, minor, spere; FIELD, group, woust; FLICK, erase, drome; FLUTE, board, proad; FRAUD, loose, joble; GAUGE, swirl, thind; GROUP, class, saind; HAZEL, proud, tound; HEDGE, cobra, froat; ICING, alert, belve; INERT, plaza, colly; KNIFE, cocoa, tould; KNOCK, badge, lible; LEVEL, knock, forst; LOCAL, hedge, ifter; LODGE, charm, thant; LOOSE, dwarf, thich; LOTUS, ridge, dince; LUNCH, pizza, brove; LYRIC, venue, shent; MIMIC, extra, plare; MINOR, black, selly; MORAL, icing, hings; MOTOR, wheel, spand; MOTTO, scrub, shise; MOURN, space, chate; NURSE, timid, mally; OBESE, fault, hally; OZONE, lyric, hught; PANDA, flute, sheme; PANEL, twist, toody; PANIC, lodge, buter; PIECE, brown, mords; PILOT, zebra, cance; PIZZA, lotus, gress; PLAZA, bison, yourn; PROUD, eagle, tathe; QUAKE, motor, forry; RAPID, motto, houst; REIGN, local, coult; RIDGE, lunch, thany; RIFLE, coach, chand; ROUGH, panic, bleas; SCOOP, rifle, mande; SCRAP, bulge, welve; SCRUB, delta, goven; SOLID, crumb, rught; SPACE, bluff, dound; SPASM, crook, iffer; SPOON, water, whild; STAFF, crude, moure; STUNT, cargo, dider; STYLE, fraud, drand; SWARM, piece, fince; SWIRL, quake, manne; SYRUP, total, theve; THEFT, moral, sping; THORN, avail, bably; THUMB, vodka, wrive; TIMID, hazel, wence; TOTAL, knife, bence; TULIP, crowd, morge; TWIST, cloud, arand; VENUE, spasm, traid; VODKA, syrup, hince; VOICE, atlas, strat; WATER, solid, shood; WHEEL, spoon, prand; ZEBRA, tulip, chout

Nonword targets

ARAND, shise, bulge; BABLY, shing, voice; BECON, dight, syrup; BELVE, prous, twist; BENCE, shist, rapid; BLEAS, rught, crowd; BROVE, wasky, atlas; BUTER, sping, scoop; CANCE, houst, thumb; CHAND, morge, swirl; CHATE, prind, bingo; CHAVE, tould, pilot; CHISE, tound, mourn; CHOUT, gress, field; CLUSE, froat, thorn; COLLY, drand, water; COULT, wasis, badge; DIDER, hally, total; DIGHT, forse, moral; DINCE, whoss, rough; DOUND, welve, alert; DRAND, chise, wheel; DROME, thich, tulip; FALLY, becon, motto; FINCE, mally, group; FORRY, bleas, lunch; FORSE, hught, timid; FORST, wence, panel; FROAT, lighe, venue; GOVEN, whild, fraud; GRESS, thant, black; GROME, thind, class; HALLY, wrive, inert; HINCE, spost, motor; HINGS, moure, extra; HOUST, iffer, lyric; HUGHT, saind, board; IFFER, yough, lotus; IFTER, chand, spasm; ITTER, dound, local;

JOBLE, traid, charm; LIBLE, thany, spoon; LIGHE, mords, dwarf; MALLY, dince, theft; MANDE, forst, flick; MANNE, thost, proud; MORDS, tathe, gauge; MORGE, quall, staff; MOURE, spand, avail; NOUND, chate, erase; PLARE, woust, stunt; PRAND, cluse, loose; PRIND, theve, hazel; PRING, thele, coach; PROAD, hince, style; PROUS, chave, eagle; QUALL, shent, ridge; QUESS, thild, minor; RIEVE, wough, blunt; ROURE, hings, plaza; RUGHT, mande, obese; SAIND, grome, booze; SELLY, prand, crook; SHEME, fally, knock; SHENT, bably, bluff; SHING, coult, fault; SHISE, toody, brown; SHIST, drome, cargo; SHOOD, fince, knife; SOULD, rieve, mimic; SPAND, ifter, flute; SPERE, chout, cocoa; SPING, forry, delta; SPOST, arand, icing; STRAT, colly, lodge; TATHE, pring, bison; THAKE, sould, crumb; THANT, lible, crude; THANY, roure, cloud; THELE, nound, scrub; THEVE, proad, solid; THICH, belve, quake; THILD, goven, space; THIND, brove, swarm; THOST, bence, reign; TOODY, manne, panic; TOULD, spere, scrap; TOUND, selly, level; TRAID, sheme, cycle; WASIS, joble, ozone; WASKY, itter, piece; WELVE, yourn, cobra; WENCE, strat, vodka; WHILD, quess, nurse; WHOSS, buter, pizza; WOUGH, dider, blitz; WOUST, cance, rifle; WRIVE, shood, panda; YOUGH, plare, zebra; YOURN, thake, hedge

B.8 Stimuli of Experiment 8

Word targets

ALERT, union, ocsym; ATLAS, power, mzdpe; AVAIL, crown, mnstw; BADGE, point, wzzyc; BINGO, speed, avkra; BISON, watch, aklph; BLACK, judge, fegyp; BLITZ, cover, cungd; BLUFF, night, apvxa; BLUNT, offer, raxri; BOARD, fight, phgil; BOOZE, thing, pujuw; BROWN, cause, sxzel; BULGE, visit, roccp; CARGO, build, qilil; CHARM, would, syilb; CLASS, begin, miqyd; CLOUD, press, miqxp; COACH, spend, fgipe; COBRA, sight, migpv; COCOA, right, lprph; CROOK, event, bvenn; CROWD, leave, sugfx; CRUDE, allow, giagj; CRUMB, plant, fdpei; CYCLE, about, aivpg; DELTA, rough, chnzk; DWARF, house, imkzo; EAGLE, round, spnfr; ERASE, pound, micpj; EXTRA, could, suvsc; FAULT, order, dcpsh; FIELD, coach, rrzsh; FLICK, under, axayw; FLUTE, broad, axypg; FRAUD, limit, voysl; GAUGE, world, syniy; GROUP, admit, eacyd; HAZEL, count, dqrki; HEDGE, plain, bxyia; ICING, short, jaxxa; INERT, laugh, uckos; KNIFE, trust, xvump; KNOCK, still, gihlw; LEVEL, brain, hymbp; LOCAL, three, ftnzi; LODGE, first, chriv; LOOSE, again, miguz; LOTUS, where, cigii; LUNCH, party, fbdfi; LYRIC, state, saqzs; MIMIC, plate, rojrp;

MINOR, equal, hfdef; MORAL, quiet, pekiq; MOTOR, class, acgph; MOTTO, large, sujcw; MOURN, table, celqi; NURSE, catch, igicw; OBESE, child, jazjh; OZONE, march, agtfw; PANDA, fully, clyym; PANEL, thick, migxu; PANIC, below, mvsym; PIECE, major, sugln; PILOT, seven, audbc; PIZZA, never, cljos; PLAZA, force, hcneo; PROUD, beach, ecajw; QUAKE, front, mihjw; RAPID, often, bfmsh; REIGN, coast, sykdd; RIDGE, youth, abwwa; RIFLE, match, hkaub; ROUGH, since, miytd; SCOOP, bread, mqava; SCRAP, light, mjymb; SCRUB, might, oyizi; SOLID, every, rakpw; SPACE, worth, rojul; SPASM, other, wzxoo; SPOON, early, cgkra; STAFF, green, ncllb; STUNT, place, ievzb; STYLE, grand, dazko; SWARM, being, dvnel; SWIRL, hence, empxn; SYRUP, whole, djekb; THEFT, brown, oljul; THORN, small, ziwzw; THUMB, local, safcn; TIMID, close, nyxoo; TOTAL, bring, imimj; TULIP, above, embxz; TWIST, clear, acnhc; VENUE, start, ricjg; VODKA, price, suuhp; VOICE, stand, auqjs; WATER, young, cgglb; WHEEL, court, icsxa; ZEBRA, sound, cxnio

Nonword targets

ARAND, syilb, whole; BABLY, hcneo, seven; BECON, gihlw, right; BELVE, spnfr, court; BENCE, miguz, light; BLEAS, dqrki, worth; BROVE, sykdd, still; BUTER, phgil, local; CANCE, migxu, sight; CHAND, fgipe, press; CHATE, voysl, union; CHAVE, imkzo, sound; CHISE, mjymb, fully; CHOUT, mvsym, being; CLUSE, miqyd, front; COLLY, abwwa, grand; COULT, empxn, where; DIDER, saqzs, coast; DIGHT, cljos, never; DINCE, suuhp, short; DOUND, bxyia, plate; DRAND, ievzb, close; DROME, axypg, again; FALLY, cxnio, trust; FINCE, lprph, start; FORRY, imimj, leave; FORSE, cgglb, match; FORST, aivpg, place; FROAT, sugln, spend; GOVEN, icsxa, thick; GRESS, fbdfi, youth; GROME, hkaub, stand; HALLY, suvsc, judge; HINCE, apvxa, party; HINGS, audbc, early; HOUST, giagj, clear; HUGHT, rakpw, plain; IFFER, mnstw, catch; IFTER, hymbp, laugh; ITTER, wzxoo, allow; JOBLE, ftnzi, first; LIBLE, sujcw, crown; LIGHE, dazko, count; MALLY, chriv, offer; MANDE, rrzsh, rough; MANNE, rojrp, child; MORDS, pekiq, equal; MORGE, acnhc, plant; MOURE, dcpsh, class; NOUND, embxz, beach; PLARE, miytd, thing; PRAND, ziwzw, quiet; PRIND, uckos, cause; PRING, ecajw, house; PROAD, celqi, event; PROUS, mihjw, admit; QUALL, fdpei, often; QUESS, jazjh, coach; RIEVE, auqjs, watch; ROURE, jaxxa, visit; RUGHT, eacyd, since; SAIND, rojul, power; SELLY, acgph, march; SHEME, agtfw, broad; SHENT, wzzyc, could; SHING, xvump, about; SHISE, pujuw, round; SHIST, cungd, major; SHOOD, raxri, table; SOULD, chnzk, three; SPAND, qilil, fight; SPERE, nyxoo,

night; SPING, djekb, other; SPOST, ncllb, every; STRAT, oljul, would; TATHE, igicw, world; THAKE, syniy, brown; THANT, mzdpe, build; THANY, roccp, order; THELE, ricjg, young; THEVE, axayw, brain; THICH, dvnel, under; THILD, sugfx, force; THIND, ocsym, above; THOST, cgkra, begin; TOODY, aklph, large; TOULD, migpv, hence; TOUND, miqxp, price; TRAID, clyym, below; WASIS, hfdef, green; WASKY, cigii, point; WELVE, oyizi, might; WENCE, mqava, small; WHILD, safcn, cover; WHOSS, avkra, bring; WOUGH, sxzel, state; WOUST, fegyp, bread; WRIVE, bfmsh, pound; YOUGH, bvenn, speed; YOURN, micpj, limit

B.9 Stimuli of Experiment 9

Word targets

WOULD, manne, hymbp; ABOUT, chise, chriv; COULD, gress, jazjh; OTHER, bably, cigii; FIRST, bence, acnhc; WHERE, colly, dazko; STILL, prand, rojrp; NEVER, sould, audbc; BEING, tould, dcpsh; RIGHT, welve, empxn; AGAIN, cluse, wzxoo; WORLD, cance, chnzk; MIGHT, belve, voysl; THREE, dound, ncllb; HOUSE, fally, cgkra; UNDER, thost, bfmsh; EVERY, shing, acgph; PLACE, forry, dqrki; THING, mords, embxz; SMALL, goven, bvenn; SINCE, toody, mqava; OFTEN, quall, hkaub; YOUNG, strat, icsxa; NIGHT, drome, ecajw; CHILD, buter, rakpw; WHOLE, arand, bxyia; POINT, sheme, mjymb; PARTY, hings, hcneo; LARGE, thind, uckos; ROUND, bleas, phgil; POWER, hally, suvsc; EARLY, hught, sujcw; STATE, prind, roccp; LIGHT, roure, axayw; ORDER, sping, qilil; FRONT, mally, saqzs; LOCAL, shent, mnstw; LEAVE, whoss, ricjg; CLEAR, shist, syniy; ABOVE, thild, suuhp; START, lible, fegyp; CLASS, itter, fbdfi; TABLE, yough, mihjw; CLOSE, traid, axypg; UNION, chate, clyym; SHORT, wence, jaxxa; MAJOR, hince, cungd; BRING, theve, eacyd; FORCE, thany, gihlw; SOUND, thele, raxri; GREEN, thich, auqjs; STAND, grome, celqi; CAUSE, dight, imimj; PRESS, whild, oyizi; BELOW, drand, giagj; COURT, saind, syilb; SEVEN, proad, miqyd; BEGIN, woust, lprph; WATCH, rieve, djekb; WORTH, fince, fgipe; COVER, wasky, miytd; OFFER, spand, aklph; VISIT, nound, rojul; PRICE, houst, oljul; SIGHT, plare, abwwa; BROWN, lighe, aivpg; FIGHT, mande, ocsym; ALLOW, rught, hfdef; SPEND, froat, migxu; QUIET, chand, cgglb; COUNT, wrive, fdpei; SPEED, chout, agtfw; BREAD, coult, miguz; BUILD, chave, rrzsh; PLANT, quess, ziwzw; FULLY, spere, micpj; CATCH, moure, sugln; BRAIN, spost, cljos; TRUST, dince, cxnio; THICK, forse, apvxa;

EQUAL, forst, imkzo; YOUTH, iffer, safcn; MARCH, tound, ievzb; BEACH, pring, ftnzi; EVENT, shood, miqxp; MATCH, selly, nyxoo; LAUGH, brove, mzdpe; JUDGE, thant, mvsym; PLAIN, morge, sugfx; ADMIT, becon, wzzyc; COAST, dider, pujuw; POUND, thake, avkra; GRAND, joble, sxzel; ROUGH, wasis, dvnel; BROAD, shise, xvump; LIMIT, wough, sykdd; PLATE, yourn, igicw; HENCE, prous, migpv; COACH, ifter, spnfr; CROWN, tathe, pekiq

Nonword targets

JVGIO, manne, hymbp; FMPAK, chise, chriv; UXWIY, gress, jazjh; HURDV, bably, cigii; SLULV, bence, acnhc; FWRWI, colly, dazko; HWEIU, prand, rojrp; JENZP, sould, audbc; MKZKA, tould, dcpsh; JOFCI, welve, empxn; AITFB, cluse, wzxoo; QGOUM, cance, chnzk; QAFQN, belve, voysl; EVAFQ, dound, ncllb; TBPIV, fally, cgkra; YNJIP, thost, bfmsh; WUTFX, shing, acgph; VUNCZ, forry, dqrki; APLVP, mords, embxz; AWYPD, goven, bvenn; SZUXS, toody, mqava; YIVTR, quall, hkaub; PHBEL, strat, icsxa; SLQTU, drome, ecajw; IMLHL, buter, rakpw; ZVCEF, arand, bxyia; KCRAO, sheme, mjymb; YQUQD, hings, hcneo; ELGPW, thind, uckos; UNJDX, bleas, phgil; EFGPE, hally, suvsc; RBFEO, hught, sujcw; AXFVB, prind, roccp; BTINM, roure, axayw; UDXRF, sping, qilil; TEJCU, mally, saqzs; ACXVG, shent, mnstw; YABYQ, whoss, ricjg; ARVWA, shist, syniy; ZEGXF, thild, suuhp; OAHJA, lible, fegyp; VASYH, itter, fbdfi; ELCQD, yough, mihjw; NEZSU, traid, axypg; RXUOO, chate, clyym; IKZGP, wence, jaxxa; ABQTB, hince, cungd; QSPIB, theve, eacyd; XZSCU, thany, gihlw; BGZOU, thele, raxri; ZWDEV, thich, auqjs; TBWNA, grome, celqi; KBESV, dight, imimj; VUMQS, whild, oyizi; UKTWY, drand, giagj; TEMXV, saind, syilb; SZKXU, proad, miqyd; XIDNI, woust, lprph; UXHSY, rieve, djekb; KAQRB, fince, fgipe; VEVRJ, wasky, miytd; XGWIE, spand, aklph; IBTXC, nound, rojul; IKBVW, houst, oljul; IQCMD, plare, abwwa; SWSNU, lighe, aivpg; PUFKI, mande, ocsym; SJSAA, rught, hfdef; ELYZJ, froat, migxu; EWSJP, chand, cgglb; ZMALZ, wrive, fdpei; KYEVD, chout, agtfw; AVFKH, coult, miguz; BFJUX, chave, rrzsh; RPLDA, quess, ziwzw; ODLAQ, spere, micpj; AYZWI, moure, sugln; YAWRF, spost, cljos; SYWUK, dince, cxnio; TLIBM, forse, apvxa; AYWPC, forst, imkzo; QHLUX, iffer, safcn; AFJKJ, tound, ievzb; AWLWH, pring, ftnzi; KVWTE, shood, miqxp; DAHFJ, selly, nyxoo; NUFLQ, brove, mzdpe; LEKRC, thant, mvsym; QBINV, morge, sugfx; FGMAF, becon, wzzyc; LKAZO, dider, pujuw; SCIBU, thake, avkra; KUPYV, joble, sxzel; FCRUJ, wasis, dvnel; FQAJF, shise, xvump; FVEXA, wough,

sykdd; SQKEF, yourn, igicw; NAXNB, prous, migpv; HGUMK, ifter, spnfr; SUXYW, tathe, pekiq

B.10 Stimuli of Experiment 10

Word targets

ALERT, nound, uckos; ATLAS, goven, chriv; AVAIL, morge, rrzsh; BADGE, forry, mvsym; BINGO, chave, hfdef; BISON, plare, avkra; BLACK, ifter, xvump; BLITZ, cance, roccp; BLUFF, proad, ocsym; BLUNT, shood, rojrp; BOARD, fince, igicw; BOOZE, thany, dqrki; BROWN, shist, celqi; BULGE, saind, icsxa; CARGO, shent, suuhp; CHARM, quess, qilil; CLASS, drome, miguz; CLOUD, sheme, bvenn; COACH, rieve, fegyp; COBRA, theve, fgipe; COCOA, iffer, mnstw; CROOK, bleas, sugln; CROWD, thele, bxyia; CRUDE, whoss, gihlw; CRUMB, thind, ftnzi; CYCLE, dight, spnfr; DELTA, sping, mihjw; DWARF, becon, oljul; EAGLE, yourn, dcpsh; ERASE, chout, cgglb; EXTRA, colly, migpv; FAULT, grome, cxnio; FIELD, prous, hkaub; FLICK, spere, pujuw; FLUTE, pring, cgkra; FRAUD, hince, sxzel; GAUGE, mords, chnzk; GROUP, chise, miytd; HAZEL, dound, miqxp; HEDGE, strat, imkzo; ICING, forse, rojul; INERT, quall, aklph; KNIFE, coult, audbc; KNOCK, mally, abwwa; LEVEL, traid, sujcw; LOCAL, prind, mzdpe; LODGE, wasis, raxri; LOOSE, prand, rakpw; LOTUS, wrive, jazjh; LUNCH, dider, dazko; LYRIC, tound, bfmsh; MIMIC, buter, lprph; MINOR, chate, wzzyc; MORAL, shing, ziwzw; MOTOR, cluse, phgil; MOTTO, belve, ecajw; MOURN, tathe, aivpg; NURSE, thich, giagj; OBESE, hught, acgph; OZONE, whild, cigii; PANDA, itter, embxz; PANEL, wough, fbdfi; PANIC, selly, voysl; PIECE, rught, mjymb; PILOT, drand, auqjs; PIZZA, houst, ncllb; PLAZA, forst, imimj; PROUD, hally, syilb; QUAKE, thost, wzxoo; RAPID, joble, nyxoo; REIGN, wasky, cljos; RIDGE, bably, hymbp; RIFLE, woust, acnhc; ROUGH, mande, eacyd; SCOOP, fally, mqava; SCRAP, lible, djekb; SCRUB, thant, agtfw; SOLID, thake, empxn; SPACE, tould, miqyd; SPASM, roure, oyizi; SPOON, lighe, axayw; STAFF, moure, clyym; STUNT, brove, micpj; STYLE, arand, cungd; SWARM, toody, fdpei; SWIRL, bence, hcneo; SYRUP, chand, jaxxa; THEFT, sould, apvxa; THORN, welve, ievzb; THUMB, gress, axypg; TIMID, yough, saqzs; TOTAL, dince, pekiq; TULIP, wence, safcn; TWIST, manne, dvnel; VENUE, froat, ricjg; VODKA, shise, sugfx; VOICE, spand, sykdd; WATER, hings, suvsc; WHEEL, spost, migxu; ZEBRA, thild, syniy

Nonword targets

ABLIL, roure, cungd; ABLUN, iffer, roccp; AFFSA, belve, dvnel; APLIA, brove, hcneo; ARLUD, chise, ocsym; AWYNS, dider, miqxp; BAKKA, wrive, oljul; BIFEL, yough, sujcw; BILIR, spost, jazjh; BLEPH, strat, migxu; BLICO, spand, pujuw; BULUM, forse, pekiq; CEENA, shist, sugfx; COCRI, thele, xvump; COMLI, prand, spnfr; CULGA, prind, mnstw; DEDIO, hally, saqzs; DULEW, sping, chriv; DULUS, becon, imkzo; ECHUR, saind, imimj; EFEBA, hings, phgil; EGESK, mally, nyxoo; EGRIS, tound, aklph; ENKIS, proad, rojul; ENRYS, hught, igicw; EUMPE, thild, oyizi; FEIRA, dound, uckos; FESPE, rught, clyym; FICIM, welve, bvenn; FILDI, prous, axypg; FLEUM, pring, ricjg; FLIKI, nound, mzdpe; FOCRO, shise, auqjs; FOICT, selly, rakpw; FULPH, grome, jaxxa; FUTEW, mords, icsxa; GEDUA, whoss, mvsym; GEDUM, wasis, acnhc; GIBIR, chout, voysl; GIBUR, cance, dazko; HUDEP, froat, giagj; HYRIS, manne, agtfw; IGISS, wence, axayw; ILDIC, spere, bfmsh; ILVIS, wough, chnzk; IVIVE, tould, cgglb; JACUM, gress, sxzel; JAZDA, itter, ncllb; JOVIC, sheme, empxn; JUDYA, morge, celqi; JULIM, tathe, acgph; KAKEM, sould, wzzyc; KARJU, lible, ievzb; KAURS, joble, hfdef; KEDDS, coult, miguz; KESUR, thich, cxnio; KLUDS, theve, migpv; KNUCH, ifter, miqyd; KOCIP, quall, sugln; KULDE, forst, syniy; LEFEB, houst, apvxa; LEFOL, drand, sykdd; LIKIR, chave, audbc; LIZIM, thake, safcn; LUSEL, chand, bxyia; MIKAK, buter, suvsc; MOKIL, quess, abwwa; NAULF, rieve, mjymb; NEFEB, wasky, rrzsh; NEZAL, woust, suuhp; OBSEP, whild, cigii; OCOMP, bleas, avkra; OUCIS, plare, djekb; OUNEM, traid, dcpsh; PESPE, yourn, fbdfi; PESUR, thany, mihjw; PLEVO, thant, cgkra; POCEF, thind, gihlw; ROSCO, dight, ziwzw; ROSUR, dince, micpj; RUILM, bence, eacyd; RUKIL, shood, fegyp; SARRO, hince, ftnzi; SAULP, goven, dqrki; SEPRI, toody, mqava; SIBOU, chate, lprph; SIFLU, mande, embxz; SKAKA, drome, qilil; SUROK, lighe, fgipe; SURSA, fince, fdpei; SYSLI, moure, ecajw; TEIRP, colly, cljos; TESEM, forry, raxri; TUECT, arand, aivpg; UNENG, fally, miytd; URVIR, bably, hymbp; VARLI, thost, wzxoo; VUKET, shing, syilb; WABIA, cluse, rojrp; ZOFLI, shent, hkaub

Appendix C: Generated items for assessing the *OT3* metric

The following list contains all 1073 five letter nonwords that were generated by the OT3 software after initialising with the BNC (*The British National Corpus*, 2001). All items scored better or equal to 1.20 SD above mean. Please note, that nonword was defined as not being listed in corpus.

withe, 2.56; nothe, 2.49; thave, 2.36; thers, 2.36; thand, 2.19; yould, 2.18; prome, 2.11; thent, 2.10; buthe, 2.05; frome, 2.05; yound, 2.03; thate, 2.02; thice, 2.00; hathe, 2.00; thery, 1.96; whave, 1.95; youre, 1.94; mothe, 1.94; frove, 1.92; whice, 1.92; thend, 1.89; gothe, 1.89; bothe, 1.89; theas, 1.88; cound, 1.88; anded, 1.87; haver, 1.87; maked, 1.86; afted, 1.86; bethe, 1.85; thime, 1.85; sithe, 1.85; thems, 1.84; comme, 1.80; youse, 1.80; haved, 1.80; mithe, 1.80; puble, 1.80; beent, 1.80; ition, 1.79; arthe, 1.79; wited, 1.79; forke, 1.79; coure, 1.78; fores, 1.78; evere, 1.78; dithe, 1.78; thist, 1.78; hould, 1.78; comed, 1.78; coned, 1.77; comen, 1.77; thise, 1.77; whand, 1.77; thess, 1.77; fored, 1.77; whime, 1.76; thein, 1.76; conly, 1.76; wille, 1.76; prout, 1.75; theat, 1.75; thead, 1.75; stion, 1.74; bould, 1.74; thare, 1.74; thern, 1.74; ances, 1.74; sught, 1.74; andis, 1.73; exple, 1.73; washe, 1.73; fithe, 1.72; hater, 1.72; whing, 1.72; suble, 1.72; pland, 1.72; ation, 1.71; conce, 1.71; thade, 1.71; thout, 1.71; ofted, 1.71; forme, 1.71; hishe, 1.71; willy, 1.71; trand, 1.71; wasse, 1.71; frout, 1.70; thene, 1.70; whout, 1.69; reare, 1.69; thown, 1.69; abled, 1.69; hight, 1.69; whise, 1.68; shere, 1.68; chave, 1.68; andes, 1.68; fould, 1.68; exame, 1.67; whown, 1.67; mance, 1.67; tothe, 1.67; yount, 1.67; saing, 1.67; buted, 1.67; althe, 1.67; thest, 1.66; anced, 1.66; ither, 1.66; tooke, 1.65; bithe, 1.65; couse, 1.65; wased, 1.64; orthe, 1.64; mence, 1.64; inces, 1.64; spere, 1.64; thert, 1.64; evers, 1.64; tione, 1.64; hiche, 1.63; woure, 1.63; beire, 1.63; grome, 1.63; suche, 1.63; taked, 1.63; lothe, 1.63; prive, 1.63; mands, 1.63; inged, 1.62; thele, 1.62; pothe, 1.62; exted, 1.62; inges, 1.62; hatte, 1.62; pries, 1.62; pards, 1.62; ander, 1.61; saide, 1.61; bable, 1.61; thens, 1.60; whate, 1.60; exper, 1.60; witte, 1.60; atere, 1.60; ithis, 1.60; stere, 1.59; prown, 1.59; thres, 1.59; facce, 1.59; exped, 1.59; beend, 1.59; calle, 1.59; ables, 1.58; rithe, 1.58; pithe, 1.58; thely, 1.58; muche, 1.58; beirs, 1.58; weent, 1.58; yeare, 1.58; thile, 1.58; forly, 1.58; thred, 1.58; behis, 1.58; smand, 1.57; lable, 1.57; forld, 1.57; anthe, 1.57; shose, 1.57;

anted, 1.57; puthe, 1.57; bight, 1.57; itted, 1.57; bects, 1.57; sione, 1.56; yoult, 1.56; inced, 1.56; parge, 1.56; somme, 1.56; sould, 1.56; thiss, 1.56; weres, 1.55; cland, 1.55; ments, 1.55; alwas, 1.55; ithat, 1.55; fiche, 1.55; whers, 1.55; thind, 1.55; aught, 1.54; overe, 1.54; youst, 1.54; ingle, 1.54; pares, 1.54; houre, 1.54; brome, 1.54; wered, 1.54; cance, 1.54; befor, 1.54; inthe, 1.54; gooke, 1.54; shers, 1.54; haded, 1.54; inted, 1.54; cally, 1.54; mosed, 1.53; anday, 1.53; somed, 1.53; hance, 1.53; somen, 1.53; gover, 1.53; hothe, 1.53; sumbe, 1.53; thase, 1.52; itive, 1.52; fache, 1.52; sible, 1.52; theit, 1.52; witer, 1.52; mathe, 1.51; looke, 1.51; forre, 1.51; bacce, 1.51; prold, 1.51; forde, 1.51; thild, 1.51; andle, 1.51; sawas, 1.51; chand, 1.51; frose, 1.51; boure, 1.50; yough, 1.50; offee, 1.50; tions, 1.50; andin, 1.50; shand, 1.50; thome, 1.50; reque, 1.50; rught, 1.50; anden, 1.50; cands, 1.49; wouse, 1.49; andly, 1.49; spers, 1.49; thold, 1.49; dence, 1.49; itere, 1.49; tooks, 1.49; atted, 1.49; beeks, 1.48; whome, 1.48; jused, 1.48; sture, 1.48; waske, 1.48; prion, 1.48; dided, 1.48; rethe, 1.48; whold, 1.48; cours, 1.48; wiled, 1.47; stive, 1.47; thowe, 1.47; twere, 1.47; reand, 1.47; thich, 1.47; expre, 1.47; mande, 1.47; warge, 1.47; whiss, 1.47; thics, 1.47; andid, 1.47; thass, 1.46; methe, 1.46; seent, 1.46; thady, 1.46; trome, 1.46; ithey, 1.46; noter, 1.46; twore, 1.46; proad, 1.46; conts, 1.46; theld, 1.46; goved, 1.46; hince, 1.46; hithe, 1.46; duche, 1.46; colve, 1.46; frold, 1.46; youle, 1.46; whind, 1.46; preas, 1.46; whowe, 1.46; succe, 1.45; thint, 1.45; aters, 1.45; thaps, 1.45; oneve, 1.45; socce, 1.45; wayst, 1.45; numbe, 1.45; lated, 1.45; sters, 1.45; rould, 1.45; mards, 1.45; pring, 1.45; ative, 1.45; outhe, 1.44; cathe, 1.44; bache, 1.44; youts, 1.44; hatin, 1.44; ineve, 1.44; foure, 1.44; tords, 1.44; cated, 1.44; resse, 1.44; proce, 1.44; shome, 1.44; haten, 1.44; isted, 1.44; gethe, 1.44; wared, 1.44; nould, 1.44; ofter, 1.44; rence, 1.44; beive, 1.44; dince, 1.43; vithe, 1.43; imple, 1.43; thims, 1.43; shold, 1.43; wasts, 1.43; thill, 1.43; offin, 1.43; nower, 1.43; sence, 1.43; wever, 1.43; sions, 1.43; nithe, 1.43; preve, 1.43; plare, 1.43; vothe, 1.43; theys, 1.43; saire, 1.42; thats, 1.42; reent, 1.42; thads, 1.42; whild, 1.42; fince, 1.42; inget, 1.42; witin, 1.42; mille, 1.42; heres, 1.42; theed, 1.42; doess, 1.42; anned, 1.42; makin, 1.42; prand, 1.42; hices, 1.41; polve, 1.41; thels, 1.41; thost, 1.41; showe, 1.41; thast, 1.41; froad, 1.41; extre, 1.41; wases, 1.41; sculd, 1.41; tople, 1.41; eques, 1.41; mucce, 1.41; youte, 1.41; worke, 1.41; heire, 1.41; conse, 1.41; hered, 1.41; drand, 1.41; brove, 1.40; thear, 1.40; winge, 1.40; worts, 1.40; saind, 1.40; moned, 1.40; coult, 1.40; spose, 1.40; aftes, 1.40; thany, 1.40; maing, 1.40; forse, 1.40; lawas, 1.40; anage, 1.40; bacts, 1.40;

facke, 1.40; havis, 1.40; wores, 1.40; beeme, 1.40; rothe, 1.40; liche, 1.40; malle, 1.40; ancle, 1.40; mords, 1.40; beene, 1.40; whost, 1.39; agand, 1.39; combe, 1.39; buter, 1.39; dishe, 1.39; somes, 1.39; expen, 1.39; hings, 1.39; nowas, 1.39; twout, 1.39; werse, 1.39; reave, 1.39; dinge, 1.39; sethe, 1.39; itime, 1.39; ingen, 1.39; froce, 1.39; monly, 1.39; coust, 1.39; offir, 1.39; faing, 1.39; saine, 1.39; ithed, 1.39; wored, 1.38; comis, 1.38; joble, 1.38; thant, 1.38; inage, 1.38; thows, 1.38; cothe, 1.38; poing, 1.38; frand, 1.38; foree, 1.38; rects, 1.38; scand, 1.38; whics, 1.38; caust, 1.38; antre, 1.38; prowe, 1.38; acked, 1.38; ittle, 1.38; finge, 1.37; pathe, 1.37; toome, 1.37; ision, 1.37; ithes, 1.37; sairs, 1.37; strat, 1.37; forne, 1.37; pated, 1.37; notte, 1.37; onere, 1.37; soned, 1.37; weend, 1.37; gooks, 1.37; itand, 1.37; prous, 1.37; wathe, 1.37; becon, 1.37; tiond, 1.37; milly, 1.37; enthe, 1.37; ented, 1.37; leare, 1.37; wated, 1.37; shice, 1.37; bouse, 1.37; safte, 1.37; yethe, 1.37; alled, 1.37; whint, 1.36; mythe, 1.36; itice, 1.36; coved, 1.36; dings, 1.36; thies, 1.36; whows, 1.36; weved, 1.36; parce, 1.36; exces, 1.36; saile, 1.36; yourn, 1.36; trive, 1.36; inere, 1.36; mared, 1.36; yeand, 1.36; wount, 1.36; minge, 1.36; werve, 1.36; fooke, 1.36; facks, 1.36; sures, 1.36; tould, 1.36; foris, 1.36; mosse, 1.35; ifted, 1.35; sonly, 1.35; andre, 1.35; folve, 1.35; beres, 1.35; shost, 1.35; torke, 1.35; trage, 1.35; bette, 1.35; fings, 1.35; ingue, 1.35; apper, 1.35; avere, 1.35; torts, 1.35; unces, 1.35; ancre, 1.35; iters, 1.35; exter, 1.35; beare, 1.35; mally, 1.35; maide, 1.35; opere, 1.35; bused, 1.35; comin, 1.34; anand, 1.34; prour, 1.34; pleve, 1.34; humbe, 1.34; sured, 1.34; hades, 1.34; whill, 1.34; ancen, 1.34; monce, 1.34; tores, 1.34; intre, 1.34; sents, 1.34; hille, 1.34; thard, 1.34; bered, 1.34; prost, 1.34; ining, 1.34; cande, 1.34; stime, 1.34; togre, 1.34; dethe, 1.34; apped, 1.34; andat, 1.34; thelf, 1.34; carge, 1.34; knote, 1.34; nowne, 1.34; toold, 1.34; hiced, 1.34; chate, 1.34; mings, 1.33; deare, 1.33; tweve, 1.33; hught, 1.33; arked, 1.33; crome, 1.33; reake, 1.33; faide, 1.33; lethe, 1.33; wites, 1.33; hande, 1.33; thoss, 1.33; parre, 1.33; nowle, 1.33; clude, 1.33; scome, 1.33; trove, 1.33; tored, 1.33; beted, 1.33; shate, 1.33; hatme, 1.33; areve, 1.33; fices, 1.33; inand, 1.33; twers, 1.33; hised, 1.33; pects, 1.33; agere, 1.33; preat, 1.33; worme, 1.33; paing, 1.33; knowe, 1.33; wayin, 1.32; surce, 1.32; thous, 1.32; ithad, 1.32; antle, 1.32; trund, 1.32; wours, 1.32; theen, 1.32; frowe, 1.32; pread, 1.32; soure, 1.32; lible, 1.32; kince, 1.32; whare, 1.32; const, 1.32; brive, 1.32; parly, 1.32; thell, 1.32; backe, 1.32; goves, 1.32; therd, 1.32; lacce, 1.32; exhis, 1.32; gived, 1.32; cithe, 1.32; lince, 1.32; iture, 1.32; frous, 1.32; proge, 1.32;

pents, 1.32; stice, 1.32; thame, 1.31; ittly, 1.31; whoss, 1.31; parld, 1.31; wasys, 1.31; butte, 1.31; yourt, 1.31; exche, 1.31; doest, 1.31; onces, 1.31; bries, 1.31; truld, 1.31; wance, 1.31; atime, 1.31; peres, 1.31; sonce, 1.31; forin, 1.31; lieve, 1.31; whous, 1.31; logre, 1.31; locce, 1.31; upple, 1.31; deres, 1.31; suffe, 1.31; spect, 1.31; morke, 1.31; becte, 1.30; hards, 1.30; tiont, 1.30; toods, 1.30; imand, 1.30; gaing, 1.30; saime, 1.30; morts, 1.30; wasee, 1.30; trame, 1.30; inger, 1.30; fouse, 1.30; chice, 1.30; onews, 1.30; maged, 1.30; aften, 1.30; reace, 1.30; prood, 1.30; peare, 1.30; thour, 1.30; coule, 1.30; twome, 1.30; prods, 1.30; moses, 1.30; ithen, 1.30; beens, 1.30; anter, 1.30; losed, 1.30; maket, 1.30; fland, 1.30; siond, 1.30; incle, 1.30; hosed, 1.30; pered, 1.30; refor, 1.30; mages, 1.30; excle, 1.30; booke, 1.30; thars, 1.30; whade, 1.30; frone, 1.29; prole, 1.29; itter, 1.29; dered, 1.29; attle, 1.29; saild, 1.29; mente, 1.29; frour, 1.29; pance, 1.29; inews, 1.29; hadis, 1.29; wrive, 1.29; doese, 1.29; throd, 1.29; nound, 1.29; crand, 1.29; reard, 1.29; aning, 1.29; onthe, 1.29; dooke, 1.29; intle, 1.29; onted, 1.29; wence, 1.29; dower, 1.29; behat, 1.29; ithas, 1.29; prits, 1.29; atice, 1.29; dight, 1.29; itake, 1.29; dould, 1.29; rathe, 1.29; ducce, 1.29; baces, 1.29; couts, 1.29; betwe, 1.29; whent, 1.28; istre, 1.28; grage, 1.28; moure, 1.28; mored, 1.28; prows, 1.28; wided, 1.28; dides, 1.28; munce, 1.28; jushe, 1.28; whour, 1.28; exced, 1.28; torms, 1.28; efuld, 1.28; havid, 1.28; wries, 1.28; foren, 1.28; asted, 1.28; bance, 1.28; nowed, 1.28; atche, 1.28; coner, 1.28; calre, 1.28; offor, 1.28; mayst, 1.28; achis, 1.28; wasto, 1.28; diche, 1.28; hader, 1.28; arche, 1.28; usted, 1.28; bence, 1.28; trave, 1.28; orked, 1.28; yeave, 1.28; kinge, 1.28; grout, 1.28; iment, 1.28; thole, 1.27; shent, 1.27; parke, 1.27; whies, 1.27; hents, 1.27; poste, 1.27; frive, 1.27; funce, 1.27; iting, 1.27; exthe, 1.27; dieve, 1.27; mants, 1.27; arged, 1.27; paide, 1.27; theye, 1.27; linge, 1.27; unthe, 1.27; taing, 1.27; agive, 1.27; pacce, 1.27; warce, 1.27; shoss, 1.27; thron, 1.27; unted, 1.27; torme, 1.27; onage, 1.27; unced, 1.27; maken, 1.27; itche, 1.27; prent, 1.27; hount, 1.27; becto, 1.27; thore, 1.27; forry, 1.27; arges, 1.27; wasks, 1.27; seare, 1.27; worce, 1.27; crive, 1.27; tathe, 1.27; tence, 1.26; histe, 1.26; weare, 1.26; aress, 1.26; forst, 1.26; shous, 1.26; antly, 1.26; nowee, 1.26; clare, 1.26; atent, 1.26; froge, 1.26; thied, 1.26; wrome, 1.26; tated, 1.26; ature, 1.26; arted, 1.26; pords, 1.26; arded, 1.26; dayst, 1.26; thart, 1.26; thict, 1.26; dathe, 1.26; thily, 1.26; alles, 1.26; preed, 1.26; fieve, 1.26; spece, 1.26; hooke, 1.26; stent, 1.26; pross, 1.26; thire, 1.26; exace, 1.26; rever, 1.26; prest, 1.26; schis, 1.26; seend, 1.26; goome, 1.26; ougge, 1.26; witle, 1.26; whins, 1.26;

herse, 1.25; nalle, 1.25; hadde, 1.25; afthe, 1.25; werst, 1.25; quire, 1.25; reate, 1.25; cleve, 1.25; lards, 1.25; walle, 1.25; werne, 1.25; greas, 1.25; cluse, 1.25; woult, 1.25; oftes, 1.25; lache, 1.25; gaide, 1.25; coute, 1.25; ancer, 1.25; moste, 1.25; rembe, 1.25; agave, 1.25; harge, 1.25; ficed, 1.25; incre, 1.25; offes, 1.25; achat, 1.25; sucts, 1.25; dowas, 1.25; frood, 1.25; parde, 1.25; teres, 1.25; pands, 1.25; frods, 1.25; affer, 1.25; thdre, 1.25; juses, 1.24; camed, 1.24; offed, 1.24; incen, 1.24; offen, 1.24; nowit, 1.24; knome, 1.24; cauld, 1.24; camen, 1.24; andit, 1.24; mushe, 1.24; frole, 1.24; arnme, 1.24; surne, 1.24; appre, 1.24; conat, 1.24; hashe, 1.24; smake, 1.24; thets, 1.24; singe, 1.24; shour, 1.24; goven, 1.24; warre, 1.24; compt, 1.24; trate, 1.24; ifere, 1.24; grame, 1.24; prese, 1.24; mater, 1.24; usion, 1.24; quile, 1.24; morms, 1.24; lould, 1.24; belve, 1.24; sques, 1.24; woust, 1.24; conot, 1.23; wasis, 1.23; tolve, 1.23; swere, 1.23; goice, 1.23; tered, 1.23; nable, 1.23; siont, 1.23; mumbe, 1.23; uning, 1.23; romme, 1.23; attly, 1.23; evend, 1.23; hists, 1.23; salle, 1.23; witen, 1.23; samed, 1.23; ithre, 1.23; loome, 1.23; maded, 1.23; onced, 1.23; inion, 1.23; bount, 1.23; youge, 1.23; torks, 1.23; frows, 1.23; quals, 1.23; samen, 1.23; forat, 1.23; theal, 1.23; thips, 1.23; hisse, 1.23; intly, 1.23; forit, 1.23; efore, 1.23; conee, 1.23; anown, 1.23; warly, 1.23; moven, 1.23; anxis, 1.23; ences, 1.23; sayst, 1.23; thain, 1.23; schat, 1.23; thood, 1.23; areat, 1.23; annis, 1.23; oners, 1.23; fathe, 1.22; morme, 1.22; clund, 1.22; holve, 1.22; greve, 1.22; herve, 1.22; roble, 1.22; wayes, 1.22; famen, 1.22; aread, 1.22; ingre, 1.22; caund, 1.22; ament, 1.22; quage, 1.22; warld, 1.22; anche, 1.22; stres, 1.22; beffe, 1.22; welve, 1.22; weves, 1.22; yeake, 1.22; posee, 1.22; socke, 1.22; colle, 1.22; cosed, 1.22; pethe, 1.22; goold, 1.22; thund, 1.22; dider, 1.22; hared, 1.22; taide, 1.22; hasse, 1.22; onand, 1.22; mosts, 1.22; quess, 1.22; aname, 1.22; iners, 1.22; iscre, 1.22; fards, 1.22; poice, 1.22; obled, 1.22; equit, 1.21; govis, 1.21; witat, 1.21; thals, 1.21; saill, 1.21; reend, 1.21; facte, 1.21; prons, 1.21; shole, 1.21; drome, 1.21; bleve, 1.21; oughe, 1.21; wayse, 1.21; wasid, 1.21; prect, 1.21; eards, 1.21; beems, 1.21; begre, 1.21; atter, 1.21; wheme, 1.21; thavy, 1.21; trion, 1.21; womme, 1.21; himed, 1.21; roure, 1.21; heare, 1.21; nathe, 1.21; reart, 1.21; bried, 1.21; proat, 1.21; torce, 1.21; shime, 1.21; tound, 1.21; himen, 1.21; heive, 1.21; speas, 1.21; inume, 1.21; whood, 1.21; prien, 1.21; spand, 1.21; rible, 1.21; stred, 1.21; butes, 1.21; nated, 1.21; toody, 1.21; firse, 1.21; baced, 1.21; ovent, 1.21; romed, 1.21; thirs, 1.21; fross, 1.21; youll, 1.21; wayed, 1.20; conte, 1.20; nally, 1.20; romen, 1.20; lacts, 1.20; avers, 1.20; movis, 1.20; pache, 1.20;

sugge, 1.20; courn, 1.20; sains, 1.20; sheme, 1.20; prody, 1.20; wally, 1.20; youry, 1.20; chere, 1.20; gence, 1.20; quirs, 1.20; mucts, 1.20; opers, 1.20; defor, 1.20; wayme, 1.20; publy, 1.20; wasin, 1.20

Appendix D: Default parameters of computational models

D.1 Bayesian Reader

Table D.1 lists the default parameters of the Bayesian Reader (Norris & Kinoshita, 2008) that were used in the simulations in Chapter 4.

Table D.1: Default parameters of the Bayesian Reader.

Parameter	Value	Parameter	Value
InitialSD	10.0	Average	50
PositionSD	5.0	LetterRanking	1
		WordRanking	10
UseLetterFrequency	off	MaxSteps	1500
SetWordPriors	on	MinSteps	5
SetLetterPriors	off	PrimeSteps	30
SetProbePriors	off	VirtualNonWordFrequency	0
UseBackgroundNonWords	off	P_a_WordThreshold	0.90 0.01 50

D.2 Spatial Coding Model

Table D.2 lists the default parameters of the Spatial Coding Model (Davis, 2010) that were used in all simulations, unless stated otherwise.

Table D.2: Default parameters of the Spatial Coding Model.

Parameter	Value	Description
dt	0.050	Step size when integrating activity equations
alpha_fl	0.280	Feature-level excitation
gamma_fl	6.000	Feature-letter inhibition
alpha_lw	0.100	Letter-word excitation
gamma_lw	0.040	Letter-word inhibition
alpha_wl	0.300	Word-letter excitation
cp	2.500	Power used for contrast-enhancing bottom up contrast
gamma_ww	0.340	Word-word inhibition
alpha_ww	0.440	Word-word self-excitation
decay_l	0.000	Letter activity decay rate

Table D.2: Default parameters of the Spatial Coding Model.

Parameter	Value	Description
decay_w	1.000	Word activity decay rate
decay_k	0.400	Match param
decay_d	5.000	Delay before decay begins
mf	0.350	Masking power
max_l	1.000	Maximum letter node activity
max_w	1.000	Maximum word node activity
min_l	-0.200	Minimum letter node activity
min_w	-0.200	Minimum word node activity
fgain	0.050	Frequency gain (controls setting of resting activities)
mu	0.680	Local activity response threshold
sigma	0.240	Variability in position code
sigma_len_slope	0.120	Gradient of sigma by length
sigma_slope	0.000	Gradient of sigma by fixation position
threshold_w	0.000	Output threshold of word nodes
lf	0.000	Length filter for length dependent lateral inhibition
leak2	0.000	Leakage across two letter channels
tdf_threshold	0.000	Activity threshold for letter nodes to send TDF
alpha_len	0.000	Pos input to nodes that match the stim length
gamma_len	0.060	Neg input to nodes that mismatch the stim length
tdm	1.000	Gate tdf signal by bottom up input
pi	1.000	Switch that controls whether matching is position invariant
y_global	0.400	Weight assigned to global activity in ldt
y3	1.000	Weight assigned to lexical identification in ldt
n_in	0.360	Input to No-channel in ldt
lambda_yn	0.060	Inhib between yes and no channels
ldt_gain	1.000	Gain control in lexical decision channels
yes_threshold	0.800	Threshold for responding Yes in ldt
no_threshold	0.800	Threshold for responding No in ldt
freq_bias	1.800	Size of constrained network
num_candidates	30.000	
letter_noise	0.000	Letter id noise parameter

Appendix E: Detailed results of the RT analyses in vincentiles

This appendix presents the detailed results of the analysis in deciles. The *z*-scores of reaction were used in the analysis and are presented in the tables as a function of the condition, i.e. incongruent/congruent or high/low prime typicality. The analysis was computed using *lmer* of the *lme4* package (Bates & Sarkar, 2007) in R (R development core team, 2007). The tables present the respective t-values and the *p*-values that were generated using a MonteCarlo Markov chain (*p*MCMC) and the standard *p*-values according to the *t*-distribution. Brysbaert (Brysbaert, 2007) noted that the *p*-values that were derived using the MCMC computation tend to be more conservative compared to the *t*-distribution.

E.1 Analysis in deciles of Experiment 1

Table E.1: Results of the analysis by decile of pronounceable nonword targets in Experiment 1 as a function of target typicality.

Decile	High typicality	Low typicality	*Effect*	*t*-value	*p*MCMC	*p* (|*t*|>0)
1	-0.86	-1.23	*0.37*	8.70	0.00	0.00
2	-0.51	-0.96	*0.45*	10.90	0.00	0.00
3	-0.29	-0.76	*0.47*	12.04	0.00	0.00
4	-0.08	-0.61	*0.53*	13.12	0.00	0.00
5	0.12	-0.47	*0.59*	14.61	0.00	0.00
6	0.40	-0.33	*0.73*	18.22	0.00	0.00
7	0.72	-0.14	*0.86*	21.68	0.00	0.00
8	1.11	0.06	*1.05*	25.83	0.00	0.00
9	1.64	0.47	*1.17*	28.86	0.00	0.00
10	2.29	1.37	*0.92*	21.72	0.00	0.00

Table E.2: Results of the analysis by decile of unpronounceable nonword targets in Experiment 1 as a function of target typicality.

Decile	High typicality	Low typicality	*Effect*	*t*-value	*p*MCMC	*p* (|*t*|>0)
1	-0.98	-1.18	*0.20*	5.21	0.00	0.00
2	-0.68	-0.94	*0.26*	7.70	0.00	0.00
3	-0.48	-0.79	*0.31*	9.00	0.00	0.00
4	-0.32	-0.66	*0.34*	9.74	0.00	0.00
5	0.01	-0.53	*0.54*	10.43	0.00	0.00
6	0.22	-0.39	*0.61*	11.76	0.00	0.00
7	0.50	-0.24	*0.74*	13.44	0.00	0.00
8	0.99	0.02	*0.97*	14.10	0.00	0.00
9	0.87	0.44	*0.43*	15.97	0.00	0.00
10	1.69	1.27	*0.42*	11.62	0.00	0.00

Table E.3: Results of the analysis by decile of low typicality nonword targets in Experiment 1 as a function of target pronounceability.

Decile	Unpro-nounceable	Pro-nounceable	*Effect*	*t*-value	*p*MCMC	*p* (|*t*|>0)
1	-1.23	-1.18	*-0.05*	-1.52	0.13	0.13
2	-0.96	-0.94	*-0.02*	-0.31	0.77	0.76
3	-0.76	-0.79	*0.03*	0.73	0.47	0.46
4	-0.61	-0.66	*0.05*	1.33	0.19	0.18
5	-0.47	-0.53	*0.06*	1.61	0.11	0.11
6	-0.33	-0.39	*0.06*	1.84	0.07	0.07
7	-0.14	-0.24	*0.10*	2.72	0.00	0.01
8	0.06	0.02	*0.04*	1.28	0.20	0.20
9	0.47	0.44	*0.03*	0.88	0.40	0.38
10	1.37	1.27	*0.10*	2.63	0.01	0.01

E.2 Analysis in deciles of Experiment 2

Table E.4: Results of the analysis by decile of word targets in Experiment 2 as a function of response congruency.

Decile	Incongruent	Congruent	*Effect*	*t*-value	*p*MCMC	*p* (|*t*|>0)
1	-1.20	-1.43	*0.23*	7.44	0.00	0.00
2	-0.86	-1.07	*0.21*	7.64	0.00	0.00
3	-0.62	-0.87	*0.25*	9.11	0.00	0.00
4	-0.43	-0.69	*0.26*	9.49	0.00	0.00
5	-0.27	-0.53	*0.26*	9.68	0.00	0.00
6	-0.10	-0.35	*0.25*	9.34	0.00	0.00
7	0.10	-0.14	*0.24*	8.88	0.00	0.00
8	0.38	0.13	*0.25*	9.25	0.00	0.00
9	0.81	0.53	*0.28*	10.22	0.00	0.00
10	1.64	1.35	*0.29*	10.27	0.00	0.00

Table E.5: Results of the analysis by decile of nonword targets in Experiment 2 as a function of response congruency.

Decile	Incongruent	Congruent	*Effect*	*t*-value	*p*MCMC	*p* (|*t*|>0)
1	-0.97	-1.22	*0.25*	8.42	0.00	0.00
2	-0.60	-0.84	*0.24*	8.85	0.00	0.00
3	-0.39	-0.63	*0.24*	8.85	0.00	0.00
4	-0.20	-0.46	*0.26*	9.19	0.00	0.00
5	-0.03	-0.28	*0.25*	9.32	0.00	0.00
6	0.15	-0.11	*0.26*	9.61	0.00	0.00
7	0.35	0.10	*0.25*	9.04	0.00	0.00
8	0.60	0.35	*0.25*	9.18	0.00	0.00
9	1.00	0.80	*0.20*	7.39	0.00	0.00
10	1.80	1.73	*0.07*	2.55	0.01	0.01

E.3 Analysis in deciles of Experiment 3

Table E.6: Results of the analysis by decile of word targets in Experiment 3 as a function of response congruency.

Decile	Incongruent	Congruent	*Effect*	*t*-value	*p*MCMC	*p* (|*t*|>0)
1	-1.26	-1.39	*0.13*	4.29	0.00	0.00
2	-0.92	-1.08	*0.16*	5.71	0.00	0.00
3	-0.73	-0.87	*0.14*	4.90	0.00	0.00
4	-0.56	-0.72	*0.16*	5.49	0.00	0.00
5	-0.39	-0.55	*0.16*	5.55	0.00	0.00
6	-0.22	-0.38	*0.16*	5.52	0.00	0.00
7	-0.02	-0.19	*0.17*	5.91	0.00	0.00
8	0.24	0.04	*0.20*	6.77	0.00	0.00
9	0.62	0.44	*0.18*	6.23	0.00	0.00
10	1.38	1.23	*0.15*	5.00	0.00	0.00

Table E.7: Results of the analysis by decile of nonword targets in Experiment 3 as a function of response congruency.

Decile	Incongruent	Congruent	*Effect*	*t*-value	*p*MCMC	*p* (|*t*|>0)
1	-0.85	-1.11	*0.26*	8.70	0.00	0.00
2	-0.54	-0.79	*0.25*	9.04	0.00	0.00
3	-0.32	-0.57	*0.25*	8.76	0.00	0.00
4	-0.16	-0.38	*0.22*	7.76	0.00	0.00
5	-0.02	-0.19	*0.17*	5.90	0.00	0.00
6	0.14	0.04	*0.10*	3.66	0.00	0.00
7	0.33	0.27	*0.06*	2.09	0.04	0.04
8	0.56	0.56	*0.00*	0.09	0.90	0.93
9	0.96	1.07	*-0.11*	-3.74	0.00	0.00
10	1.87	1.95	*-0.08*	-2.49	0.01	0.01

E.4 Analysis in deciles of Experiment 4

Table E.8: Results of the analysis by decile of word targets in Experiment 4 as a function of response congruency.

Decile	Incongruent	Congruent	*Effect*	*t*-value	*p*MCMC	*p* (|*t*|>0)
1	-1.13	1.27	*-2.40*	4.62	0.00	0.00
2	-0.83	-0.98	*0.15*	5.46	0.00	0.00
3	-0.62	-0.79	*0.17*	6.13	0.00	0.00
4	-0.44	-0.62	*0.18*	6.31	0.00	0.00
5	-0.28	-0.42	*0.14*	4.94	0.00	0.00
6	-0.09	-0.23	*0.14*	5.06	0.00	0.00
7	0.12	-0.01	*0.13*	4.72	0.00	0.00
8	0.36	0.28	*0.08*	2.97	0.00	0.00
9	0.78	0.73	*0.05*	2.05	0.05	0.04
10	1.60	1.55	*0.05*	1.83	0.07	0.07

Table E.9: Results of the analysis by decile of nonword targets in Experiment 4 as a function of response congruency.

Decile	Incongruent	Congruent	*Effect*	*t*-value	*p*MCMC	*p* (|*t*|>0)
1	-1.00	-1.21	*0.21*	6.98	0.00	0.00
2	-0.70	-0.84	*0.14*	5.13	0.00	0.00
3	-0.50	-0.65	*0.15*	5.25	0.00	0.00
4	-0.35	-0.49	*0.14*	5.04	0.00	0.00
5	-0.20	-0.34	*0.14*	4.94	0.00	0.00
6	-0.04	-0.17	*0.13*	4.69	0.00	0.00
7	0.16	0.04	*0.12*	4.41	0.00	0.00
8	0.43	0.33	*0.10*	3.68	0.00	0.00
9	0.84	0.76	*0.08*	3.00	0.00	0.00
10	1.61	1.71	*-0.10*	-3.27	0.00	0.00

E.5 Analysis in deciles of Experiment 5

Table E.10: Results of the analysis by decile of word targets in Experiment 5 as a function of response congruency.

Decile	Incongruent	Congruent	*Effect*	*t*-value	*p*MCMC	*p* (|*t*|>0)
1	-1.29	-1.40	*0.11*	2.68	0.00	0.01
2	-1.00	-1.09	*0.09*	2.49	0.01	0.01
3	-0.79	-0.87	*0.08*	2.43	0.02	0.02
4	-0.63	-0.71	*0.08*	2.41	0.02	0.02
5	-0.45	-0.56	*0.11*	2.99	0.01	0.00
6	-0.26	-0.37	*0.11*	3.29	0.00	0.00
7	-0.05	-0.15	*0.10*	2.80	0.01	0.01
8	0.25	0.18	*0.07*	2.23	0.03	0.03
9	0.69	0.64	*0.05*	1.46	0.17	0.15
10	1.49	1.45	*0.04*	0.92	0.36	0.36

Table E.11: Results of the analysis by decile of nonword targets in Experiment 5 as a function of response congruency.

Decile	Incongruent	Congruent	*Effect*	*t*-value	*p*MCMC	*p* (|*t*|>0)
1	-0.92	-1.05	*0.13*	3.52	0.00	0.00
2	-0.61	-0.72	*0.11*	3.28	0.00	0.00
3	-0.41	-0.50	*0.09*	2.90	0.00	0.00
4	-0.23	-0.29	*0.06*	2.04	0.04	0.04
5	-0.07	-0.12	*0.05*	1.69	0.10	0.09
6	0.11	0.05	*0.06*	1.74	0.08	0.08
7	0.34	0.25	*0.09*	2.77	0.01	0.01
8	0.62	0.52	*0.10*	3.17	0.00	0.00
9	1.02	0.94	*0.08*	2.58	0.01	0.01
10	1.90	1.73	*0.17*	4.88	0.00	0.00

E.6 Analysis in deciles of Experiment 6

Table E.12: Results of the analysis by decile of word targets in Experiment 6 as a function of response congruency.

Decile	Incongruent	Congruent	*Effect*	*t*-value	*p*MCMC	*p* (|*t*|>0)
1	-1.29	-1.32	*0.03*	1.08	0.31	0.28
2	-1.00	-1.06	*0.06*	1.95	0.05	0.05
3	-0.80	-0.85	*0.05*	1.93	0.07	0.05
4	-0.61	-0.71	*0.10*	2.70	0.01	0.01
5	-0.45	-0.54	*0.09*	3.08	0.00	0.00
6	-0.25	-0.35	*0.10*	2.82	0.00	0.00
7	-0.02	-0.14	*0.12*	3.80	0.00	0.00
8	0.25	0.13	*0.12*	3.17	0.00	0.00
9	0.67	0.55	*0.12*	3.72	0.00	0.00
10	1.52	1.30	*0.22*	6.62	0.00	0.00

Table E.13: Results of the analysis by decile of nonword targets in Experiment 6 as a function of response congruency.

Decile	Incongruent	Congruent	*Effect*	*t*-value	*p*MCMC	*p* (|*t*|>0)
1	-0.97	-1.05	*0.08*	2.36	0.02	0.02
2	-0.67	-0.72	*0.05*	1.63	0.10	0.10
3	-0.48	-0.52	*0.04*	1.16	0.27	0.25
4	-0.32	-0.33	*0.01*	0.60	0.55	0.55
5	-0.13	-0.16	*0.03*	1.02	0.31	0.31
6	0.05	0.05	*0.00*	0.06	0.97	0.96
7	0.27	0.30	*-0.03*	-1.20	0.23	0.23
8	0.55	0.59	*-0.04*	-1.38	0.15	0.17
9	1.01	1.02	*-0.01*	-0.54	0.59	0.59
10	1.78	1.85	*-0.07*	-2.15	0.03	0.03

E.7 Analysis in deciles of Experiment 7

Table E.14: Results of the analysis by decile of word targets in Experiment 7 as a function of response congruency.

Decile	Incongruent	Congruent	*Effect*	*t*-value	*p*MCMC	*p* (|*t*|>0)
1	-1.21	-1.21	*0.00*	-0.17	0.84	0.87
2	-1.01	-1.01	*0.00*	0.28	0.80	0.78
3	-0.88	-0.88	*0.00*	0.02	0.99	0.98
4	-0.76	-0.77	*0.01*	0.20	0.85	0.84
5	-0.62	-0.65	*0.03*	0.92	0.34	0.36
6	-0.49	-0.51	*0.02*	0.56	0.58	0.57
7	-0.32	-0.33	*0.01*	0.64	0.51	0.52
8	-0.08	-0.10	*0.02*	0.73	0.46	0.47
9	0.30	0.32	*-0.02*	-0.41	0.71	0.68
10	1.14	1.22	*-0.08*	-2.27	0.02	0.02

Table E.15: Results of the analysis by decile of nonword targets in Experiment 7 as a function of response congruency.

Decile	Incongruent	Congruent	*Effect*	*t*-value	*p*MCMC	*p* (|*t*|>0)
1	-0.73	-0.75	*0.02*	0.62	0.53	0.53
2	-0.47	-0.50	*0.03*	0.93	0.37	0.35
3	-0.29	-0.31	*0.02*	0.91	0.37	0.36
4	-0.10	-0.15	*0.05*	1.32	0.18	0.19
5	0.06	0.05	*0.01*	0.54	0.59	0.59
6	0.24	0.24	*0.00*	-0.06	0.93	0.95
7	0.47	0.49	*-0.02*	-0.51	0.63	0.61
8	0.79	0.74	*0.05*	1.74	0.08	0.08
9	1.19	1.20	*-0.01*	-0.24	0.80	0.81
10	1.99	1.99	*0.00*	-0.03	0.99	0.97

E.8 Analysis in deciles of Experiment 8

Table E.16: Results of the analysis by decile of word targets in Experiment 8 as a function of response congruency.

Decile	Incongruent	Congruent	*Effect*	*t*-value	*p*MCMC	*p* (|*t*|>0)
1	-1.25	-1.30	*0.05*	1.49	0.14	0.14
2	-1.00	-1.60	*0.60*	1.78	0.07	0.07
3	-0.85	-0.91	*0.06*	2.03	0.04	0.04
4	-0.71	-0.78	*0.07*	2.26	0.02	0.02
5	-0.59	-0.66	*0.07*	2.46	0.01	0.01
6	-0.44	-0.52	*0.08*	2.99	0.00	0.00
7	-0.26	-0.36	*0.10*	3.07	0.00	0.00
8	-0.03	-0.11	*0.08*	2.63	0.01	0.01
9	0.38	0.28	*0.10*	3.61	0.00	0.00
10	1.29	1.18	*0.11*	3.47	0.00	0.00

Table E.17: Results of the analysis by decile of nonword targets in Experiment 8 as a function of response congruency.

Decile	Incongruent	Congruent	*Effect*	*t*-value	*p*MCMC	*p* (|*t*|>0)
1	-0.81	-0.78	*-0.03*	-0.95	0.36	0.34
2	-0.49	-0.49	*0.00*	-0.21	0.87	0.84
3	-0.31	-0.28	*-0.03*	-0.77	0.41	0.44
4	-0.14	-0.12	*-0.02*	-0.80	0.43	0.43
5	0.01	0.05	*-0.04*	-1.39	0.17	0.17
6	0.19	0.23	*-0.04*	-1.62	0.10	0.11
7	0.39	0.45	*-0.06*	-2.01	0.05	0.05
8	0.64	0.77	*-0.13*	-4.60	0.00	0.00
9	1.10	1.19	*-0.09*	-3.13	0.00	0.00
10	1.91	1.99	*-0.08*	-2.70	0.01	0.01

E.9 Analysis in deciles of Experiment 9

Table E.18: Results of the analysis by decile of word targets in Experiment 9 as a function of prime typicality.

Decile	Low typicality	High typicality	*Effect*	*t*-value	*p*MCMC	*p* (|*t*|>0)
1	-1.31	-1.43	*0.12*	3.52	0.00	0.00
2	-0.95	-1.13	*0.18*	5.48	0.00	0.00
3	-0.71	-0.91	*0.20*	6.02	0.00	0.00
4	-0.51	-0.72	*0.21*	6.31	0.00	0.00
5	-0.33	-0.56	*0.23*	6.96	0.00	0.00
6	-0.11	-0.37	*0.26*	7.89	0.00	0.00
7	0.11	-0.17	*0.28*	8.76	0.00	0.00
8	0.37	0.11	*0.26*	7.91	0.00	0.00
9	0.79	0.53	*0.26*	8.13	0.00	0.00
10	1.56	1.43	*0.13*	3.71	0.00	0.00

Table E.19: Results of the analysis by decile of nonword targets in Experiment 9 as a function of prime typicality.

Decile	Low typicality	High typicality	*Effect*	*t*-value	*p*MCMC	*p* (|*t*|>0)
1	-1.25	-0.93	*-0.32*	-9.65	0.00	0.00
2	-0.87	-0.59	*-0.28*	-9.17	0.00	0.00
3	-0.64	-0.38	*-0.26*	-8.27	0.00	0.00
4	-0.42	-0.19	*-0.23*	-7.31	0.00	0.00
5	-0.22	0.00	*-0.22*	-7.07	0.00	0.00
6	0.00	0.17	*-0.17*	-5.43	0.00	0.00
7	0.22	0.35	*-0.13*	-3.98	0.00	0.00
8	0.54	0.59	*-0.05*	-1.61	0.10	0.11
9	0.91	0.97	*-0.06*	-1.93	0.05	0.05
10	1.77	1.76	*0.01*	0.25	0.82	0.81

E.10 Analysis in deciles of Experiment 10

Table E.20: Results of the analysis by decile of word targets in Experiment 10 as a function of prime typicality.

Decile	Low typicality	High typicality	*Effect*	*t*-value	*p*MCMC	*p* (|*t*|>0)
1	-1.15	-1.22	*0.07*	2.16	0.03	0.03
2	-0.89	-0.95	*0.06*	2.16	0.02	0.03
3	-0.72	-0.77	*0.05*	1.96	0.05	0.05
4	-0.56	-0.60	*0.04*	1.68	0.09	0.09
5	-0.41	-0.47	*0.06*	1.57	0.11	0.12
6	-0.24	-0.27	*0.03*	0.88	0.38	0.38
7	-0.04	-0.08	*0.04*	1.00	0.32	0.32
8	0.27	0.20	*0.07*	1.61	0.11	0.11
9	0.72	0.66	*0.06*	0.79	0.44	0.43
10	1.57	1.64	*-0.07*	-4.70	0.00	0.00

Table E.21: Results of the analysis by decile of nonword targets in Experiment 10 as a function of prime typicality.

Decile	Low typicality	High typicality	*Effect*	*t*-value	*p*MCMC	*p* (|*t*|>0)
1	-1.04	-0.97	*-0.07*	-2.54	0.01	0.01
2	-0.73	-0.70	*-0.03*	-1.19	0.24	0.24
3	-0.55	-0.51	*-0.04*	-1.49	0.14	0.14
4	-0.40	-0.35	*-0.05*	-1.99	0.05	0.05
5	-0.26	-0.21	*-0.05*	-1.90	0.06	0.06
6	-0.07	-0.06	*-0.01*	-0.44	0.65	0.66
7	0.14	0.13	*0.01*	0.35	0.72	0.73
8	0.39	0.41	*-0.02*	-0.61	0.53	0.54
9	0.87	0.87	*0.00*	0.05	0.98	0.96
10	1.78	1.79	*-0.01*	0.69	0.50	0.49