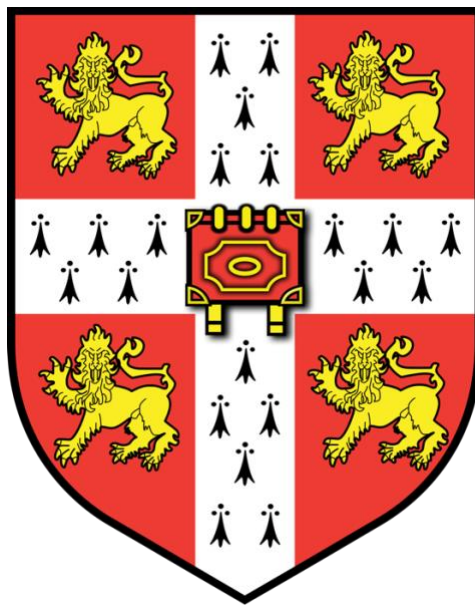


Mechanisms of Spoken Word Recognition and Memory Encoding Studied through Competitor Priming

Yingcan Wang

This thesis is submitted for the degree of
Doctor of Philosophy



Darwin College

MRC Cognition and Brain Sciences Unit

University of Cambridge

December 2021

PREFACE

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the preface and specified in the text. It is not substantially the same as any work that has already been submitted before for any degree or other qualification except as declared in the preface and specified in the text. It does not exceed the prescribed word limit of 60,000 words as specified by the Degree Committee of the Faculties of Clinical Medicine and Clinical Veterinary Medicine.

The research outlined in Chapter 2, Chapter 3 and Chapter 4 was conducted in collaboration with Ediz Sohoglu and Rebecca Gilbert, who helped in data collection and preliminary analysis.

The study reported in Chapter 2 has been published as:

Wang, Y. C., Sohoglu, E., Gilbert, R. A., Henson, R. N., & Davis, M. H. (2021). Predictive Neural Computations Support Spoken Word Recognition: Evidence from MEG and Competitor Priming. *Journal of Neuroscience*, 41(32), 6919-6932.

ABSTRACT

MECHANISMS OF SPOKEN WORD RECOGNITION AND MEMORY ENCODING STUDIED THROUGH COMPETITOR PRIMING

Yingcan Wang

Human listeners achieve quick and effortless speech comprehension in daily life and can adopt new words easily into their vocabulary. However, the underlying mechanisms under spoken word recognition and learning remain to be better understood. This thesis examines the neural and functional mechanisms of spoken word recognition and memory encoding by using a competitor priming paradigm - prior presentation (priming) of a competitor spoken word (e.g. *hijack*) is followed by the presentation of a similar sounding word sharing the same initial segments (e.g. *hygiene*). Consistent with the Bayes rule, the prior probability of the competitor word has been increased due to the earlier exposure, which can in turn change the perception or memory encoding of the target word.

The MEG study described in Chapter 2 examined the neural implementations of spoken word recognition by testing two distinct implementations of Bayes perceptual inference. Competitive-selection accounts (e.g. TRACE) propose direct competition between lexical units such that inhibition of irrelevant candidates leads to selection of critical words, while predictive-selection accounts (e.g. Predictive Coding) suggest that computations of prediction error by comparing heard and predicted speech sounds drive the update of lexical probabilities that are crucial to word recognition. The study results indicated that MEG signals localised to the superior temporal gyrus (STG) showed greater neural responses evoked by competitor primed words than unprimed words *after* the point at which they were uniquely identified (after /haidʒ/ in *hygiene*) and these stronger neural signals also correlated with the longer response

times caused by competitor priming. These findings were more in line with the predictive neural mechanisms.

Chapter 3 reports studies that investigated lexical and sub-lexical processing during spoken word recognition, specifically whether changes in lexical prediction that give rise to the competitor priming effect (longer response times) continue to be observed even when word recognition is not required for task performance. Here, the pause detection task was compared with the lexical decision task in a set of experiments to direct participants' attention to phonological processing or lexical processing respectively during the perception of prime or target items. The findings showed opposite effects of these two kinds of processing, with the competitor priming effect observable only when participants' attention was on lexical processing, while phonological facilitatory effect was observed when the pause detection task was used, and that prime item was presented with pause inserted. These results were in accordance with the Predictive Coding account and the Distributed Cohort Model, as both of which support inhibitory lexical processing and facilitatory sub-lexical processing in their respective structures.

Chapter 4 describes tasks and analyses that examined the effect of competitor priming on spoken word memory encoding by using additional recognition memory data collected from the same experiments as reported in Chapter 2 and 3. Participants' memory performance was measured by how accurately they could distinguish previously heard items from foils. The findings indicated that enhanced prediction error caused by competitor priming facilitated memory encoding of words when the encoding was repeated multiple times and involved deeper lexical-semantic processing. These findings were consistent with the PIMMS framework, which proposes that prediction error caused by the competitor priming effect should improve memory encoding. Moreover, subsequent memory analyses of the MEG data (as reported in Chapter 2) showed pseudoword encoding localised to the medial temporal lobe,

consistent with the initial rapid encoding stage of novel word learning in the complementary learning systems.

In conclusion, the thesis provides evidence for a unified account of computations of prediction error which supports spoken word recognition and memory encoding while also shows that the effects of lexical and sub-lexical processing are dissociated during these two processes.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Matt Davis. His guidance, advice, support and encouragement over the past few years have been extremely valuable to me. I am very grateful and consider myself incredibly lucky to have someone as kind, knowledgeable and professional as Matt as my supervisor and mentor.

I am also thankful for my advisor, Rik Henson, for his help and insight on my memory work and my MEG study and for his always being there even when my questions are naïve.

I would also like to thank Ed Sohoglu and Becky Gilbert. They taught me all the technical nitty-gritties involved in scientific research from the scratch. I have gained more confidence in myself thanks to their help.

I am also grateful for all the friends I have met since I arrived in the UK. They have been the source of my joy. I would like to thank my friends and colleagues at the MRC Cognition and Brain Sciences Unit (CBU), especially my office mates, who turned this place into a home that I always want to go back to. Special thanks also go to the Darwin College Boat Club, no other group can make me wake up at 5 in the morning with a smile on my face. To Shego, who has always been there supporting me. To Tongtong, for every meal we had together in Cambridge. To Xiao, my perfect rowing and interpreting partner. To Panpan, for her warm accompany.

Most importantly, I am extremely grateful for my family. To my parents, their unconditional love and support have made me who I am today. To my husband, Dan, I can't describe how happy I am with him in my life.

Finally, I would like to thank the China Scholarship Council for generously funding my PhD studentship and the CBU for financially supporting the last three months of my PhD.

TABLE OF CONTENTS

PREFACE.....	3
ABSTRACT.....	5
ACKNOWLEDGEMENTS	9
TABLE OF CONTENTS	11
1 INTRODUCTION AND LITERATURE REVIEW.....	17
1.1 The competitor priming paradigm.....	17
1.2 Functional mechanisms of spoken word recognition.....	20
1.2.1 Lexical activation.....	21
1.2.2 Probability modulation – the frequency effect.....	23
1.2.3 Lexical selection	31
1.3 Neural implementations of spoken word recognition.....	34
1.3.1 The location	35
1.3.2 The time course.....	39
1.4 Mechanisms of word memory encoding	43
1.4.1 Prediction and memory	43
1.4.2 Levels of processing and attentional resources.....	47
1.4.3 The complementary learning systems for novel word encoding	49
1.5 Thesis structure.....	54

2	PREDICTIVE NEURAL COMPUTATIONS SUPPORT SPOKEN WORD	
	RECOGNITION: EVIDENCE FROM MEG AND COMPETITOR PRIMING	57
2.1	Introduction.....	57
2.2	Methods.....	62
2.2.1	Participants.....	62
2.2.2	Experimental Design.....	62
2.2.3	Stimuli.....	64
2.2.4	Post-test Gating Study.....	67
2.2.5	Behavioural Data Analyses	70
2.2.6	MEG Data Acquisition, Processing and Analyses.....	71
2.2.7	Source Reconstruction	72
2.3	Results	74
2.3.1	Behaviour	74
2.3.2	MEG.....	76
2.4	Discussion.....	86
2.4.1	Competitive- vs predictive-selection	86
2.4.2	How do listeners process pseudowords?.....	89
2.4.3	Conclusion	90
3	LEXICAL AND SUB-LEXICAL PROCESSING IN SPOKEN WORD	
	RECOGNITION	93
3.1	Introduction.....	93
3.2	Experiment 2	99

3.2.1 Method	99
3.2.2 Results.....	102
3.2.3 Discussion	105
3.3 Experiment 3	105
3.3.1 Methods.....	106
3.3.2 Results.....	108
3.3.3 Discussion	110
3.4 Experiment 4-5	111
3.4.1 Methods.....	114
3.4.2 Results of Experiment 4a (prime: pause detection; target: lexical decision)	118
3.4.3 Results of Experiment 4b (prime: lexical decision; target: lexical decision)	123
3.4.4 Results of Experiment 5a (prime: pause detection; target: pause detection)	125
3.4.5 Results of Experiment 5b (prime: lexical decision; target: pause detection).....	129
3.4.6 Comparison of Word Targets RT results from Experiment 4 & 5	131
3.5 General Discussion.....	133
3.5.1 Competitor priming effect limited to lexical processing	133
3.5.2 Lexical and sub-lexical processing during spoken word recognition	135
3.5.3 Facilitatory effect from pseudoword priming	137
3.5.4 Conclusion	138
4 SPOKEN WORD ENCODING – ANALYSES ON MEMORY DATA FROM EXPERIMENTS 1 - 5.....	139
4.1 Introduction.....	139

4.2 Memory Data Analyses (Experiment 1-5)	146
4.2.1 Methods.....	148
4.2.2 Results & Discussion	152
Experiment 1	152
Experiment 2	156
Experiment 3	158
Experiment 4-5.....	161
4.3 Subsequent Memory Analyses of MEG Data	167
4.3.1 Methods.....	167
4.3.2 Results & Discussion	169
4.4 General discussion	176
4.4.1 Competitor priming effects modulate memory encoding	177
4.4.2 The neural mechanisms that support pseudoword encoding	181
4.4.3 Conclusion	182
5 CONCLUDING REMARKS	185
5.1 Summary.....	185
5.2 Predictive coding for spoken word recognition and learning.....	190
5.3 Lexical and sub-lexical processing for spoken word recognition and learning	192
5.4 Conclusion	193
APPENDIX A LIST OF STIMULI.....	196
APPENDIX B SUPPLEMENTARY MATERIALS FOR CHAPTER 2	212

APPENDIX C	SUPPLEMENTARY MATERIALS FOR CHAPTER 3.....	217
APPENDIX D	SUPPLEMENTARY MATERIALS FOR CHAPTER 4.....	220
BIBLIOGRAPHY	225

1 INTRODUCTION AND LITERATURE REVIEW

Speech comprehension is a fundamental human cognitive skill that affects our daily life. In order to understand speech, it is necessary that we can recognise each individual spoken word that is familiar to us and also constantly learn new words that we have not encountered before. This thesis investigates the mechanisms of spoken word recognition and memory encoding by using a competitor priming paradigm (Monsell & Hirsh, 1998).

Speech perception is a process that constantly matches segments of speech input with word forms and meanings that we know. This process unfolds over time and the speech input is often transiently ambiguous, since a speech segment (e.g. /kæptɪ/) can match multiple word candidates (e.g. *captain*, *captive*) at a given time. Therefore, in order to select the suitable match, the process of spoken word recognition is essentially competitive and the competitor priming paradigm reflects exactly such competitive process.

In this chapter, I first introduce what the competitor priming paradigm is and why it is helpful for studying spoken word recognition and memory encoding while reviewing existing studies that used this paradigm. Then, I review the theories and evidence on the functional and neural mechanisms of how spoken word is recognised and subsequently remembered and how the mechanisms involved in these processes are revealed through the competitor priming paradigm. Finally, I describe the remaining structure of the thesis.

1.1 The competitor priming paradigm

It has been known that repetition priming can enhance response speed or accuracy towards previously exposed stimuli even when considerable time has elapsed between successive presentations (Forbach et al. 1974). Such effects are prevalent during spoken word recognition

(e.g. Orfanidou et al., 2006, 2011) and are commonly used to measure implicit memory of spoken words (Church & Schacter, 1994; Goldinger, 1996). In contrast to repetition priming, competitor priming leads to an inhibitory priming effect which shows that a prior presentation of a competitor word (e.g. *hijack*) delays the response of similar sounding word that shares the same initial sounds (e.g. *hygiene*).

While this inhibitory effect caused by competitor priming was predicted by major models of spoken word recognition such as TRACE (McClelland & Elman, 1986), the cohort model (Marslen-Wilson, 1987, 1990) and Shortlist (Norris, 1994), surprisingly, the effect was unstable and even facilitatory when the competitor prime was presented very shortly before the target (Radeau et al., 1995; Slowiaczek & Pisoni, 1986; Slowiaczek & Hamburger, 1992) due to possible response biases (Goldinger et al., 1992; Norris et al., 2002). For example, Norris and colleagues (2002) conducted experiments that examined facilitation effects on the perception of spoken word targets when they rhymed with the prime items (e.g. *ramp-lamp*). They found that participants could notice the rhyming relationship between the prime and the target, which hence helped them adopt a strategy that biased their response towards “yes”. Goldinger et al. (1992) also showed similar findings that phonological primes (e.g. *ram-run*) had facilitatory effects for both words and pseudowords, which was also likely due to the fact that participants noticed the presence of related pairs. In addition, Dufour and Peereeman (2003) showed that the inhibitory effect in competitor priming with a short lag was reliable only with certain restrictions, such as target words having few competitors within the same cohort.

In order to explore the possibility of a clear and stable inhibitory effect in the competitor priming paradigm, Monsell and Hirsh (1998) conducted a series of experiments in which words and word-sounding novel words (referred to as pseudowords hereafter) having the same heads (e.g. *difficult, diffident*) or rhymes (e.g. *sharp, harp*) served as the prime and target items with either short lags (a few trials) or long lags (separate blocks of trials, 1-5 min) in between, and

participants' lexical decision performance on these items were measured. They found that response times increased for words preceded by another word sharing its head but not its rhyme when there was a long lag between them, while similar effects were not observed for words primed by pseudowords or pseudowords primed by items of either lexicality. When the prime lag was short, however, general phonological facilitatory effects and response bias were found as in previous literature (Goldinger et al., 1992; Norris et al., 2002). Dufour and Nguyen (2017) replicated these results by also using a lexical decision task with long priming lag. They also showed that the competitor priming effect remained robust when the prime and the target were presented in different voices and that the prime was presented only once.

These reliable findings on long-lag competitor priming have important implications for lexical access, competition and selection during spoken word recognition. The fact that the inhibitory priming effect was shown only for words sharing the same initial segments but not the rhyme indicated that the overlap of the initial sounds but not rhyme or overall phonological similarity defines the effective competitor during lexical activation, which is in accordance with the cohort theories (the Cohort Model, Marslen-Wilson, 1987, 1990; the Distributed Cohort Model, Gaskell & Marslen-Wilson, 1997, see section 1.2 for more details of these models). Meanwhile, just one prior presentation of the competitor word could already affect the recognition of the target word - this showed the strong modulation effect of prior probability of a word, namely the frequency effect (Dahan et al., 2001). Finally, target words were recognised more slowly due to the prior presentation of the competitor prime word while similar effects were not found for when pseudowords served as primes or targets. This could be explained by competitions between pre-existing lexical candidate through lateral inhibition (TRACE; McClelland & Elman, 1986, see section 1.2, 1.3 for more details of this model) or calculations of prediction error during lexical selection (Predictive Coding; Davis & Sohoglu, 2020, see section 1.2, 1.3 for more details of this model). These processes involved in the

competitor priming paradigm makes it an ideal platform to study both the functional and neural mechanisms of spoken word recognition.

Moreover, while word memory can be enhanced after repetition priming, it is possible that competitor priming may also improve episodic memory encoding of unexpected target items due to computations of prediction error (PIMMS account; Henson & Gagnepain, 2010, see section 1.4 for more details of this model). However, while a number of studies on recognition memory of words and visual episodic memory (Corley et al., 2007; Federmeir et al., 2007; Greve et al., 2017) showed support for the positive encoding effect of prediction error, other studies (Bernhard et al., 2009; Hölte et al. 2019) suggested that more expected words are remembered more successfully in subsequent memory tests. The competitor priming paradigm is therefore suitable for further investigating the role of prediction error in memory encoding of words and pseudowords.

Overall, both functional and neural mechanisms of spoken word recognition and word episodic memory encoding can be studied using the competitor priming paradigm. I will review these mechanisms in more detail in the following sections.

1.2 Functional mechanisms of spoken word recognition

Speech comprehension appears to be a common and easy task for most people, yet it in fact involves a challenging computational process for listeners. In daily conversation, listeners need to identify approximately 200 words/minute (Tauroza & Allison, 1990) from a vocabulary of ~40,000 words (Brysbaert et al., 2016). This means that they must recognise 3-4 words/second, constantly select from sets of transiently ambiguous words that start with the same sounds (e.g. *hijack* and *hygiene* both begin with /haidʒ/) and cannot easily go backwards to check what they have missed due to the temporal dimension of the speech signal – it unfolds over time.

In essence, the process of spoken word recognition can be described as matching the spoken input with mental representations of possible word candidates and ultimately select one that is consistent with the input. The functions of spoken word recognition were classified into three stages by Frauenfelder and Tyler (1987) – the initial contact that activate lexical representations, the selection of activated lexical candidates that match the sensory input, and the integration of selected word items with context. I will review the functions in the first two stages that are involved in competitor priming.

1.2.1 Lexical activation

It has been agreed that the recognition of spoken words is not simply phoneme perception, rather, it involves lexical activation to achieve the ease and rapidity of recognition. That being said, theories differ on how lexicons of spoken words are activated. The original and revised Cohort Models (Marslen-Wilson & Welsh, 1978; Marslen-Wilson, 1987) and the Distributed Cohort Model (DCM; Gaskell & Marslen-Wilson, 1997) emphasize on the temporal dimension of speech and suggest that words having the same onsets as the spoken word (i.e. the word-initial cohort) are activated first and, as the signal input coming in, items that no longer match the input will decrease in activation. While the original Cohort Model has strict match constraint for word-initial cohort, the revised Cohort Model and the DCM evaluate the similarity between lexical representations and speech input continuously so that the possibility that a word could be recognised even when the initial sounds are mispronounced can be accounted for (Cole, 1973). Similarly, the TRACE model (McClelland & Elman, 1986) also predicts that words sharing the same initial sounds with the input have an advantage in early activation, but it also supports activation of rhyme-overlapping candidates at a later stage of the speech input (Allopenna et al., 1998).

In contrast, the Neighbourhood Activation Model (NAM; Luce & Pisoni, 1998) differs from cohort theories by predicting that spoken words are activated by their overall similarity with the sensory input instead of giving priority to the match of initial sounds. The model proposes that words differ from the spoken word input by no more than one phoneme are activated regardless of the position of the phoneme (e.g. words such as *fat*, *cut*, *cab*, *at*, *scat* should be activated on hearing *cat*, while words such as *captain* will not be activated due to the difference in more than one phonemes).

There is empirical evidence for each of these predictions. For the NAM model, Luce & Pisoni (1998) conducted three experiments using perceptual identification, lexical decision and single-word shadowing paradigms respectively and showed that the accuracy and processing speed of spoken words were affected by their overall similarity neighbourhood. For the Cohort Model, Marslen-Wilson (1993) described results showing that when a word (e.g. *beetle*) had the same onset as the prime spoken word (e.g. *beaker*), the priming effect on its semantic associates (e.g. *insect*) was reliable in the visual lexical decision task, while this was not the case when the word had the same rhyme as the prime (e.g. *speaker*, with *stereo* as the semantic associate). In addition, Allopenna and colleagues (1998) found evidence for the TRACE model through an eye tracking experiment. They showed that participants' eye movements were fixed more on the onset-overlapping competitor earlier although there was also a small proportion of fixation on the rhyme-overlapping competitor later. Finally, the competitor priming effect observed by Mosell and Hirsh (1998) showed reliably delayed recognition only for words sharing the onset but not the rhyme, which highlighted the importance of the word-initial cohort in lexical activation.

1.2.2 Probability modulation – the frequency effect

Apart from phonological similarity to the sensory input, frequency, or the prior probability of words and sound sequences (i.e. how probable for the word to occur based on Bayes theorem, see Norris & McQueen, 2008), also play an important role in shaping both the activation and selection of lexical candidates. In this section, I review the frequency effect on word recognition from both the time dimension (i.e. the locus of the lexical frequency effect during word identification) and the functional dimension (i.e. the phonotactic probability at the sub-lexical level of processing as opposed to the lexical frequency effect).

The time locus of lexical frequency effect

While there is a consensus that word frequency influences word recognition, models diverge in when its effect comes into place. In localist models (i.e. models with discrete lexical representations), such as the Cohort Model (Marslen-Wilson, 1987) and TRACE (McClelland & Elman, 1986), frequency operates at an early stage by affecting the baseline activation level of each lexical unit or the connection weights between sublexical and lexical units. Similarly, in the Distributed Cohort Model (DCM; Gaskell & Marslen-Wilson, 1997), frequency affects the weights of the distributed lexical representations and this effect persists throughout the duration of word processing. In contrast, the Neighbourhood Activation Model (NAM; Luce & Pisoni, 1998) places the effect of frequency later after the lexical activation stage, this is because, in NAM, frequency affects the selection of words after sensory information for lexicons has been accumulated, although the procedure still happens during the recognition process.

However, Balota and Chumbley (1984) proposed an even later influence of frequency, which operates at the decision stage evoked by different tasks. They found that the word frequency effect on visual word recognition is largest in the lexical decision task, but smaller

in the pronunciation task and smallest in the category verification task even though lexical access must be involved in both lexical decision and category verification. They argued that the lexical decision task differs from the other tasks due to its emphasis on frequency information at the decision stage of the task, as the familiarity of the stimulus and its meaningfulness are the two most obvious pieces of information that help the participants differentiate words from pseudowords. It should also be noted that this account only intended to show the effect of frequency caused by decisions made during the task and did not intend to negate the frequency effect during lexical access in the early stage. However, similar effects caused by tasks were not observed by Monsell et al. (1989), their categorization and naming tasks showed similar sensitivity to frequency as the lexical decision task and they argued that it was the global monitoring of lexical activation during word identification, which the decision process depended on, that was sensitive to frequency. However, these results were based on visual word perception, the effect during spoken word recognition may differ due to the time-course by which words are heard, since late decision effects will only occur later on once words can be identified.

Empirical evidence was found for early as well as late effects of frequency. Connine and colleagues (1993) used a phoneme identification task in which participants were asked to identify the ambiguous phoneme in words (e.g. the phoneme /b/ or /p/ on a scale between *best* and *pest*), and they manipulated the “context” of the word by presenting a list of other words with different frequencies at the same time. They reported that presenting a low frequency word list biased the ambiguous word towards the low frequency choice, i.e. a reversed frequency effect, while high frequency word list exaggerated the frequency effect by biasing the ambiguous word towards the high frequency choice even more. Furthermore, they found that fast responses were affected by the list manipulation rather than by the frequency of the word

itself. Hence, they suggested that word frequency operates at the decision stage, in accordance with Balota & Chumbley's (1984) account.

In contrast, Cleland and colleagues (2006) observed an early frequency effect through the psychological refractory period (PRP) techniques, in which there is a central bottleneck that prevents processing of two sequential stimuli being carried out simultaneously. They found that the frequency effect was reliably smaller when the interval between the two stimuli was short (100ms) than when it was longer (800ms), which indicated that at least some of the frequency-sensitive processing (i.e. the extra recognition cost carried by low-frequency relative to high frequency words) occurred in the early pre-bottleneck period.

The early frequency effects were further confirmed by an ERP study conducted by Dufour and colleagues (2013). They used French words with high and low frequency to test participants' ERP response during a lexical decision task. They showed that low frequency words generated a stronger P350 (a component reflecting activation of lexical form) than high frequency words during lexical decision as early as 350ms post stimuli onset when the stimuli were still ambiguous.

Furthermore, Dahan and colleagues (2001) monitored participants' eye movements as they were instructed to click on and move pictured objects using the computer mouse. They showed in their first experiment that participants were more likely to fixate on high frequency competitor (e.g. a picture of *bed*) of the target picture (e.g. picture of a *bench*) than the low frequency one (e.g. picture of a *bell*) from as early as 267ms after the onset of the target word in the spoken instruction (e.g. "Pick up the bench."), while the magnitude of the effect grew as more of the word was heard. This means that the frequency effect occurred at a very early time of spoken word perception when the speech signal was still ambiguous and that the effect was continuous. In addition, the authors also simulated three different implementations of frequency in the TRACE model using the word stimuli from their experiments. The three

implementations respectively were frequency operating on resting activation levels, frequency operating on connection weights and frequency based on a post-activation decision rule. As shown in Figure 1.1, they found that resting level and post-activation bias produced very similar predictions, while connection weights between phonemes and words provided the most accurate predictions out of the three implementations, which was consistent with an account of continuous frequency effect which depends on the strength of the bottom-up input.

In short, previous evidence seemed to show that the frequency effect during word recognition can occur very early during the lexical activation stage, but the effect is somewhat continuous throughout the recognition of a spoken word and could also be affected by tasks at the decision level.

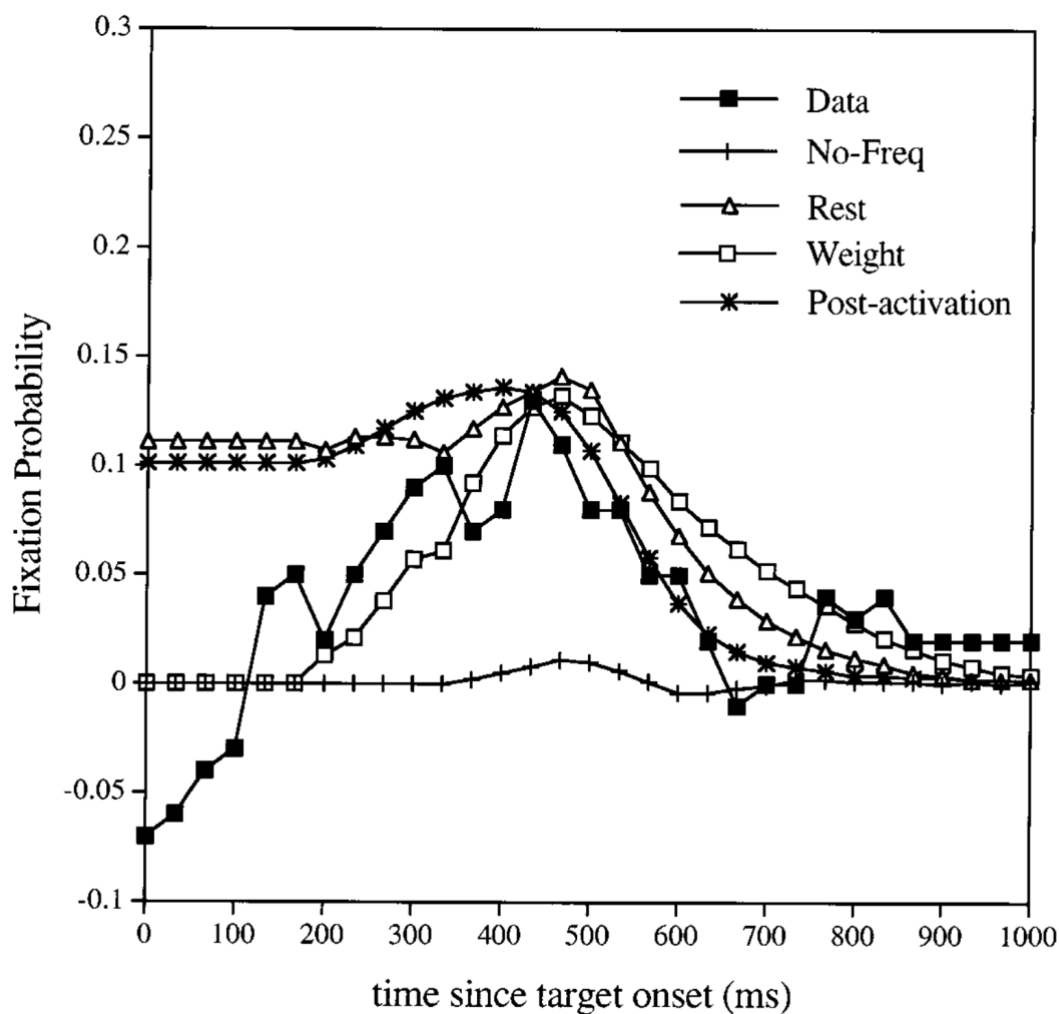


Figure 1.1. The difference of fixation probability between high and low-frequency competitors as a function of time (from Dahan et al., 2001). The actual data together with four simulated implementations of frequency were shown: simulation without frequency, simulation with frequency operating on resting activation levels, simulation with frequency operating on connection weights and simulation with frequency based on a post-activation decision rule.

Frequency effect at the sub-lexical level

Most theories agree that the process of spoken word recognition consists of both lexical and sub-lexical representations (e.g. TRACE, Shortlist, DCM). While the effect of word frequency at the lexical level has been studied widely, the effect of sub-lexical phonotactic probability, i.e. the frequency of segments or sound sequences within words, is less well known.

The TRACE model (McClelland & Elman, 1986, see Figure 1.2) consists of the feature, phoneme and word levels, hence it supports both lexical and sub-lexical level processing. However, the effects of phoneme processing are shown through excitatory interactions between the phoneme and the word layers, while word identification is driven by lateral inhibitions (more details see section 1.2.3) between lexical competitors at the word level, hence the effects of sub-lexical units on spoken word recognition are indirect and have been masked by the lexical level.

However, in the DCM (Gaskell & Marslen-Wilson, 1997, see Figure 1.3), feature input passes through the hidden layer onto the phonological and semantic representations of spoken words at the same time. The lexical and sub-lexical representations are blended together in the same nodes, hence the effect of phonotactic probabilities can be shown directly through coherent phonological representations from words within the same cohort (Gaskell & Marslen-Wilson, 1999, 2002).

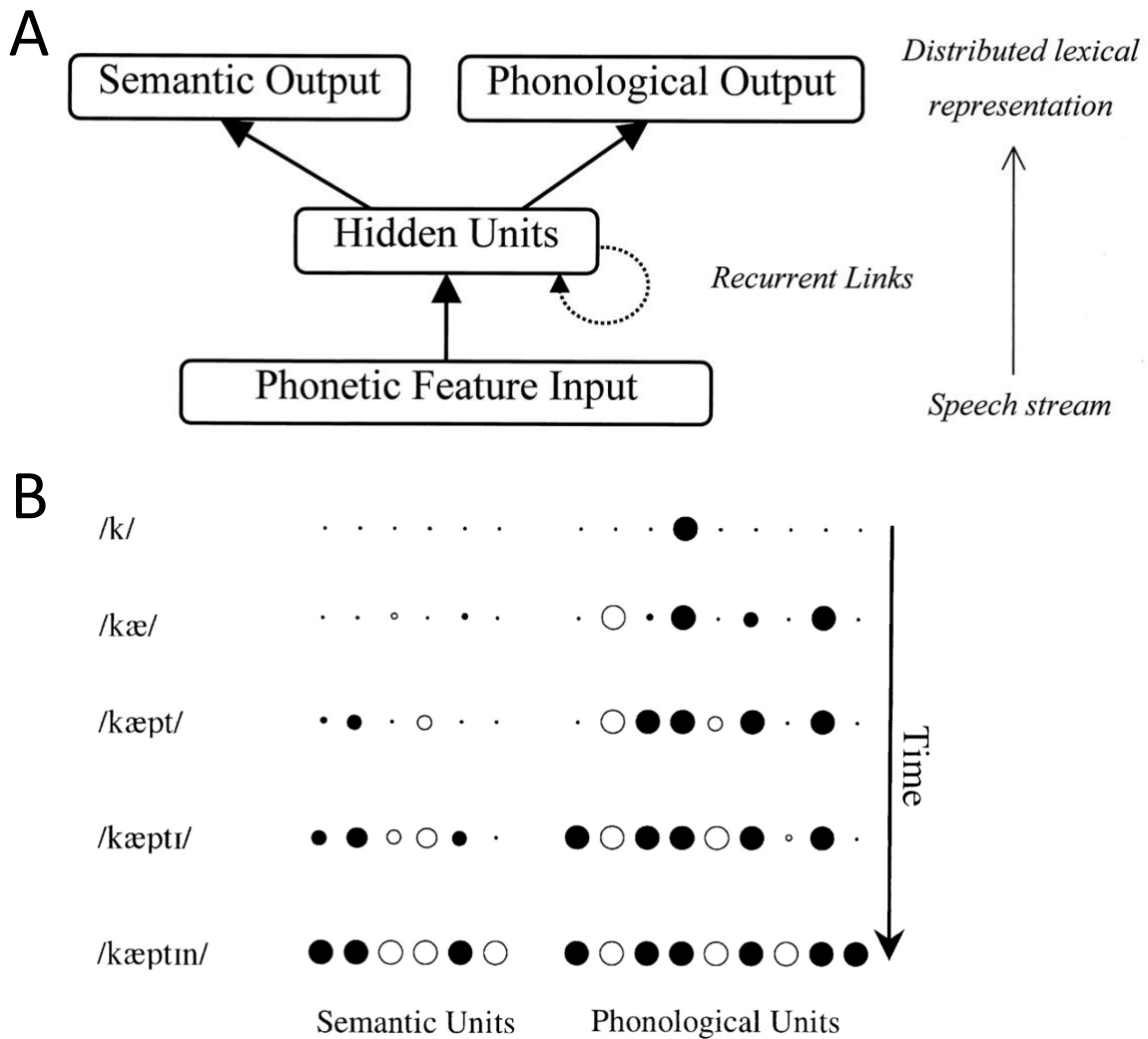


Figure 1.3. Schematic illustration of the Distributed Cohort Model (adapted from Gaskell & Marslen-Wilson, 1999, 2002). **A.** The structure of the model which shows feature input passing through the hidden units layer onto two sets of output units that represent the semantic and phonological aspects of the spoken words. **B.** Hypothetical activation pattern of semantic and phonological units as a function of time. Filled circles represent high activations while open circles represent low activations, larger circles indicate more extreme values. The activation pattern is a blend of all distributed word representations at each time step. The phonological pattern is already informative early on, since all matching candidates share the same phonological features at word onsets.

Perhaps the most detailed account of phonotactic probabilities was proposed by Vitevitch and Luce (1998, 1999, 2005). They suggested that word forms with high phonotactic probability tend to have many phonological neighbours (e.g. *cat* is high in phonotactic probability because /k/, /æ/ and /t/ are high-frequency segments in the initial, medial and final positions respectively), and that higher phonotactic probability should facilitate word recognition. However, Vitevitch et al. (1999) also showed that the sum of the frequency of segments in words positively correlated with the words' neighbourhood density, which is supposed to have an inhibitory effect on their identification. In light of this, Vitevitch and Luce (1998) proposed a two-level framework of processing in spoken word recognition, in which lexical and sub-lexical levels of representations have reversed effects. This is similar to the DCM account, which proposes intrinsic competitions between lexical representations and facilitations caused by phonological coherence. However, models such as TRACE also embody suitable architectures for this kind of dual-level processing, although it is unclear if these models as currently instantiated can simulate the opposite effects of the two levels.

An obvious example of these effects were shown through responses to words and nonwords with manipulated neighbourhood density / phonotactic probability. In a series of experiments using the shadowing task, Vitevitch and Luce (1998; 1999) showed that words with densely populated neighbourhoods were shadowed less quickly than words with sparsely populated neighbourhoods, while nonwords that were high in segment frequency were responded more quickly than those with low phonotactic probability. Additionally, the use of different experimental paradigms can also trigger similarly opposite effects from lexical and sub-lexical representations. Gaskell and Marslen-Wilson (2002) conducted a series of studies using cross model repetition priming and semantic priming paradigms respectively, with participants performing the lexical decision task. They showed that short and incomplete

spoken word segments (e.g. /kæpt/) had a significant facilitatory repetition priming effect on visual words that shared the same segments (e.g. *captain*) despite the segments' high ambiguity, whereas during the semantic priming, in which word segments primed semantic synonyms of target words (e.g. *commander* as the synonym for *captain*), there was no reliable priming effect for words with late uniqueness point (UP; i.e. the diverging point at which words can be uniquely identified). Their results indicated facilitatory effects between coherent phonological representations and competition effects between lexical semantic representations. In particular, this facilitatory phonotactic effect was caused by activations of multiple candidates who share the same initial segments, and repetition priming has enhanced the prior probability of these sound segments, which is consistent with the effect of higher phonotactic probability.

Overall, both the time and the processing level at which frequency operates can affect spoken word recognition. By presenting the competitor prime item prior to the target item, the competitor priming paradigm provides a way to manipulate the prior probability (frequency) of the prime item. While Monsell and Hirsh (1998) used the lexical decision task in their study, it remains unclear if similarly delayed word recognition effect can be found when a task not focusing on lexical processing is used. We will investigate this topic more in Chapter 3 of the thesis.

1.2.3 Lexical selection

Once word candidates are activated based on phonological similarity with the spoken input and their prior probability (i.e. frequency of occurrence), a mechanism is needed to select the matching word item in order to achieve word recognition. A number of spoken word recognition theories assume some form of competition during the lexical selection process (Cohort, DCM, NAM, Shortlist, TRACE), while the calculation of prediction error has also been proposed as a possible selection mechanism (Predictive Coding; Davis & Sohoglu, 2020).

Competitive selection

Multiple models support that spoken word recognition needs competition, but they differ in how the competition mechanism is implemented. The revised Cohort Model (Marslen-Wilson, 1987) proposes that the activation level of a unit was monitored with respect to all other units, hence the recognition of a given word item was affected by other matching candidates and the selected item fits the input the best.

In the Distributed Cohort Model (DCM; Gaskell & Marslen-Wilson, 1997), lexical competition is an intrinsic process for lexical selection, since all representations of semantic and phonological features are blended together and there is interference between patterns that are consistent with the partial input. However, partial activation and blending of multiple phonological representations will produce a lesser form of competition – since the set of partially active words all include a set of segments that are shared between lexical competitors, hence the representation of these onset segments can be readily identified.

The NAM account (Luce & Pisoni, 1998) states that the probability of recognising a word is proportional to the ratio between frequency of the target word and the total frequency of its neighbourhood, hence it also embodies competition between lexical candidates as specified in the concept of the neighbourhood.

Finally, hierarchical connectionist models such as TRACE (McClelland & Elman, 1986) and Shortlist (Norris, 1994) assume direct competition between lexical units via an explicit competition mechanism - lateral inhibition, namely a lexical item with units that are more activated within the same lexical layer (and phoneme layer, in the case of TRACE) will also directly inhibit other candidates more, which eventually leads to its selection. In addition, the recurrent loops created by lateral inhibition also allow for continuous updates of activations of candidates over time.

Predictive selection

Like TRACE and Shortlist, the Predictive Coding account (PC; Davis & Sohoglu, 2020) also supports a hierarchical structure for speech processing, with acoustic, phoneme and semantic levels. It differs from competitive selection models by incorporating inhibitory top-down predictions based on activated lexical units (see Figure 1.4 for a comparison between TRACE and the PC). This theory proposes that the strength of predictions depends on the prior probabilities of connected lexical units. By subtracting top-down predictions from bottom-up sensory signals, the resulting prediction errors at intermediate levels should drive recognition of spoken words. When sensory signals are stronger than predicted, the relevant perceptual hypotheses are strengthened and leading to positive prediction errors, while weaker signals would lead to negative prediction errors which suppress incorrect perceptual hypotheses.

Although competitive- and predictive-selection models propose different implementations for the lexical selection stage of spoken word recognition, they both approximate conditional probability in activating lexical units and their predictions for behavioural responses are similar (Spratling, 2008). For example, in the competitor priming paradigm, both TRACE and PC predict that words primed by another competitor word with the same initial segments should be recognised more slowly. From the perspective of TRACE, this is because prior presentation of the competitor word should activate it more due to the change of its frequency (Dahan et al., 2001), and hence there is stronger lateral inhibition coming from the competitor word, which makes it more difficult to select the target word. However, according to PC, prior presentation of the competitor word increases its prior probability (frequency), which hence strengthens the predictions between its initial segments and ending segments. Therefore, a larger prediction error should be evoked when this prediction is disrupted by the presentation of the target word which has a different ending

segment from the competitor word. In order to dissociate the competitive- and predictive-selection accounts, it is necessary to review evidence from neural data, which I will discuss more in the next section.

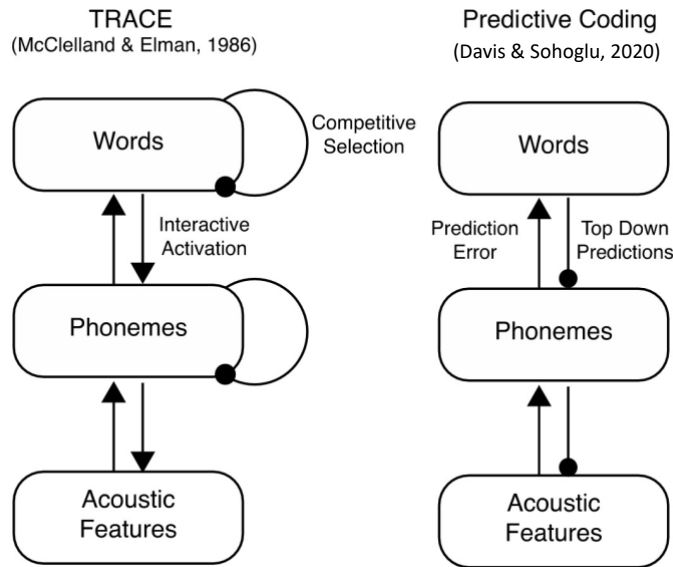


Figure 1.4. A comparison of the TRACE model (McClelland & Elman, 1986) and the Predictive Coding model (PC; Davis & Sohoglu, 2020). TRACE supports lexical selection through direct competitions (i.e. lateral inhibitions) between activated units, while PC proposes that computations of prediction error by subtracting top-down predictions from bottom-up sensory input drive the selection of the matching word (adapted from Davis & Sohoglu, 2020).

1.3 Neural implementations of spoken word recognition

Understanding the neural basis of how humans identify spoken words is helpful for us to explain the remarkable speed and flexibility of spoken word recognition. In this section, I will review the brain regions and the time course at which lexical selection happens and how competitive- and predictive-selection accounts differ in their predictions for neural responses.

1.3.1 The location

To distinguish between the brain regions that contribute to word recognition, studies contrasting neural responses to familiar words and novel words (pseudowords) are especially insightful. Davis and Gaskell (2009) reviewed 11 PET and fMRI studies, through which they observed differential responses between spoken words and pseudowords in both directions – greater word responses were shown in the inferior frontal region, several areas of the lateral temporal lobe and the temporoparietal junction, while greater pseudoword responses were mainly shown in the superior temporal gyrus (see Figure 1.5).

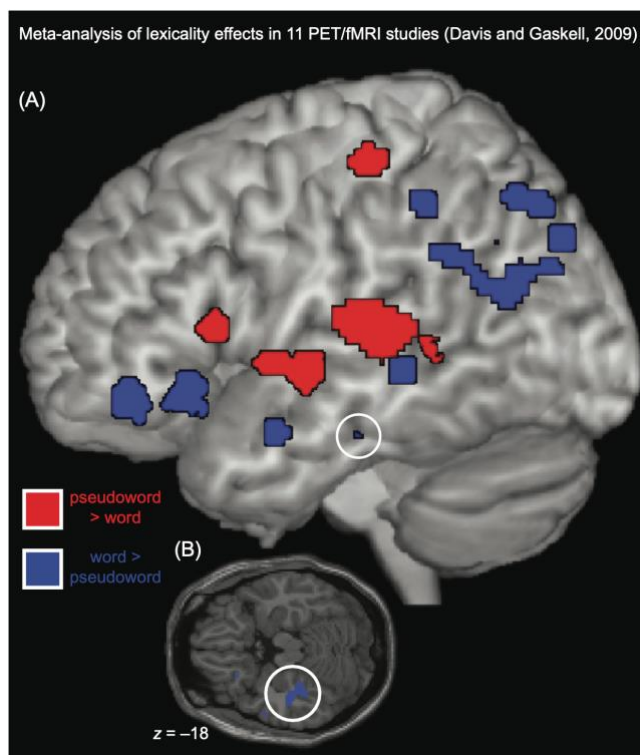


Figure 1.5. Results of a meta-analysis of 11 PET and fMRI studies comparing neural responses evoked by spoken words and pseudowords (Davis & Gaskell, 2009; Figure from Davis, 2016).

A. The Activation Likelihood Estimation maps are thresholded at $p < .05$ FDR corrected and only clusters larger than 100 mm^3 are displayed. Areas in red (mainly peri-auditory regions of

the STG) show greater neural responses evoked by pseudowords than words, areas in blue (IFG, anterior, posterior and inferior regions of the lateral temporal lobe and adjacent parietal regions) show greater word responses than pseudowords. **B.** A cross section illustration ($z = -18$) of greater word responses in inferior temporal area.

Competitive-selection models with localist structure (e.g. TRACE; McClelland & Elman, 1986) propose that lexical nodes represent words rather than pseudowords, hence they predict additional activation for words than pseudowords during lexical processing. However, it is hard for these models to explain the reverse observation of greater neural responses evoked by pseudowords in the superior temporal gyrus (STG). While the DCM (Gaskell & Marslen-Wilson, 1997) is also a competitive-selection model, it assumes common phonological representations for both words and pseudowords, hence stronger pseudoword responses could also be explained by greater processing effort engaged by sub-lexical processing in pseudowords due to their unfamiliarity (Taylor et al. 2013). The predictive-selection accounts (e.g. Predictive Coding; Davis & Sohoglu, 2020), however, predict stronger pseudoword responses than word responses in the STG due to the maximal prediction error evoked by pseudowords (since the difference between lexical informed predictions and sensory input for unfamiliar pseudowords is large).

Neural evidence exists for both competitive- and predictive-selection accounts. Additional brain activities evoked by spoken words due to lexical competition are commonly reflected in the superior temporal region (Hickok & Poeppel, 2007) and the inferior frontal cortex (Bozic et al., 2010; Zhuang et al., 2014). Okada and Hickok (2006) showed in their fMRI study that spoken words with high neighbourhood density (i.e. more lexical competition due to many similar sounding neighbours) produced significantly stronger neural responses in

the posterior superior temporal sulcus compared to low-density words. In addition, Bozic and colleagues (2010) reported an fMRI study which showed that comparison between spoken words presented with and without an embedded stem (e.g. *claim(clay)*) highlighted that bilateral inferior frontal cortex with a focus on BA47 (pars orbitalis) and BA45 (pars triangularis) were activated due to on-line lexical competition. Zhuang and colleagues (2014) also found activations of similar inferior frontal regions for words with more initially activated competitors.

In contrast, the additional responses to spoken pseudowords in the STG have been attributed to greater processing effort possibly evoked by larger prediction error in identifying pseudowords (Davis, 2016). For instance, Davis et al. (2009) and Gagnepain et al. (2012) reported similar interaction effects of neural signals between the lexicality of items and the day on which the items were presented using fMRI and MEG respectively. As shown in Figure 1.6, Gagnepain and colleagues (2012) found that pseudowords (i.e. novel words) produced stronger MEG responses than words (i.e. source words) in the left STG. Importantly, at the same location, they also found that pseudowords presented on the second day (newer, e.g. *mushrood*) evoked stronger responses than pseudowords presented on the first day (older, e.g. *formubo*), while words presented on the second day (e.g. *mushroom*) evoked weaker responses than those presented on the first day (e.g. *formula*).

They reasoned that pseudoword presented on the second day (e.g. *mushrood*) was newer and therefore evoked larger prediction error than the pseudowords presented on the first day (e.g. *formubo*, which should have been lexicalised and included in participants' vocabulary after overnight consolidation), while words presented on the first day (e.g. *formula*) should produce larger prediction error since these words share the same initial segments with the lexicalised pseudowords (e.g. *formubo*) and this newly added competitor should enlarge the difference between prior expectations (now predicting two possible segments on hearing *formu*,

i.e. “-la” and “-bo”) and the sensory input (e.g. *formula*). The interaction effects localised to the STG reflected the greater processing difficulty caused by pseudowords compared to words, newer pseudowords compared to lexicalised pseudowords, and words with more competitors compared to words with less competitors. Taken together, these findings are more in line with the predictive-selection account. In addition, it should be noted that the time-locked measures of neural activity adopted in this study also added an important time dimension that accounts for spoken word recognition. I will discuss this in more detail in the next section.

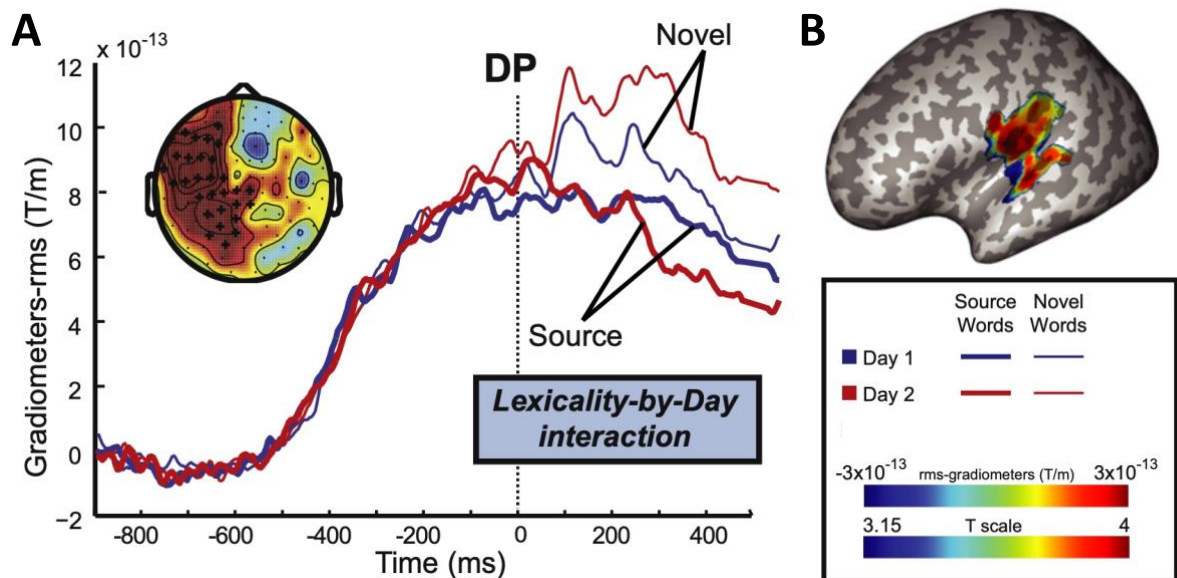


Figure 1.6. The lexicity by day interaction effect localised in the STG (adapted from Gagnepain et al., 2012). Novel words (pseudowords) on average evoked stronger MEG responses than source words. In addition, novel words presented on the second day produced greater neural responses than those presented on the first day, while source words showed more responses on the first day compared to the second day.

1.3.2 The time course

As introduced above, an important feature of speech is that it unfolds over time, it is therefore crucial to also study the neural mechanisms of spoken word recognition reflected through the time dimension. According to competitive-selection accounts, lexical competition reflects the degree of uncertainty before word identification. This can be quantified by entropy over lexical competitors activated at the point of the speech input, as described in information theory (Shannon, 1948; see Figure 1.7). Therefore, lexical entropy should be greatest when the number of competitors are the largest (i.e. at word onset), and it should be gradually reduced to minimum when a word item can be uniquely identified. While based on predictive-selection accounts, prediction error (or surprisal) is also the largest on hearing an item due to the large difference between the prediction of multiple lexically activated units and just one unit of sensory input. The prediction error should become smaller with more information input received. Note that prediction error and surprisal are similar neural measures in that prediction error is equivalent to the sum of all surprisal values associated with the heard segment. Even though the strength of lexical entropy and prediction error seem correlated in their general trend over time (Gwilliams & Davis, 2021, see Figure 1.7), they are still distinguishable with suitable manipulation.

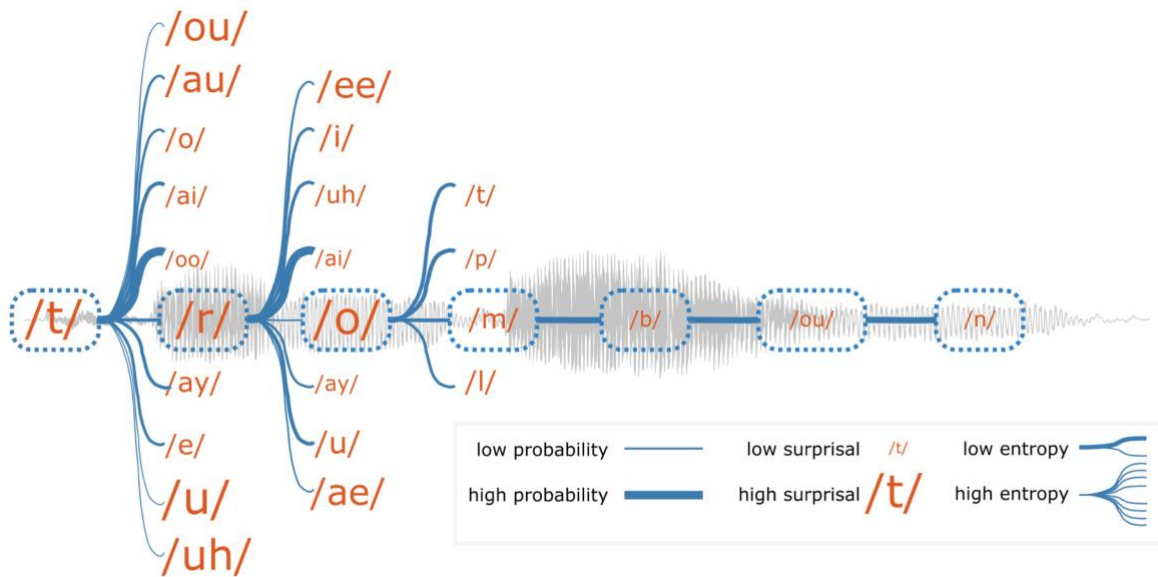


Figure 1.7. Illustration of information theory metrics measuring phoneme activations during spoken word recognition (adapted from Gwilliams & Davis, 2021). The word *trombone* is used as an example here. The blue trajectories represents the likelihood of each phoneme on hearing the previous phoneme, with the thicker ones showing higher likelihood. Entropy is represented by the number of possible continuations following each phoneme input, while the size of the phoneme (in orange) indicates the surprisal (similar to prediction error).

The MEG study by Gagnepain et al. (2012), as described in the previous section (see Figure 1.6), time-locked the neural responses to the divergence point of their stimuli (i.e. *formu* for *formula*, *formubo*). They predicted that when hearing the word *formula* on the second day of the experiment, lexical entropy before the divergence point should be enhanced, since a new competitor (*formubo*) has been added to the lexicon. Conversely, prediction error should be increased after the divergence point, since the ending segment “-la” is less expected now that “-bo” is also possible following the presentation of the initial segment *formu*. Their results showed that the MEG signal strength pattern before and after the divergence point matched with the hypotheses of predictive-selection accounts better, with reduced neural responses

before the divergence point and greater neural responses evoked 280-350ms after the divergence point for words like *formula*.

Kocagoncu and colleagues (2017) observed neural evidence for continuous lexical competition as quantified by entropy before the point at which words can be uniquely identified (UP) using the MEG. By using the LexComp model (which was defined as the change in cohort entropy from the onset until the UP), they found that early transient effects reflecting phonetic computations were shown in left STG and SMG from -400ms to -376ms before the UP, which mapped onto later effects engaging competitive processes between word candidates in left MTG and IFG around -200ms before the UP. They concluded that the location and timing of these effects were consistent with competitive-selection accounts.

A few recent studies using natural continuous speech stimuli provided neural evidence for both competitive- and predictive- selection accounts. Brodbeck and colleagues (2018) used predictor variables such as cohort size, cohort reduction (the number of items removed from cohort on hearing speech input), phoneme surprisal (i.e. an information theory term for prediction error) and cohort entropy to model MEG-recorded neural responses to continuous speech. They reported that both phoneme surprisal and entropy significantly accounted for brain responses, with predictive processing occurring at 114ms and lexical competition at 125ms after phoneme onset in the STG. In addition, Donhauser and Baillet (2019) trained a neural network which used context to predict speech phonemes and showed that early theta (4-10 Hz) responses at 60-120ms and 230ms were accounted for by lexical entropy while later delta responses (0.5-4 Hz) at 80-160ms and 230-420ms were modulated by phoneme surprisal. However, it should be noted that lexical entropy and phoneme surprisal are ordinarily correlated in continuous speech and hence hard to separate.

Overall, the responses to words in continuous speech are faster than those to words studied in isolation (Gwilliams & Davis, 2021; see Figure 1.8). It could be that the context in

natural speech has facilitated processing of individual word items and hence neural responses were evoked earlier. It may also be possible that words are produced faster when spoken in connected speech than in isolation. Experiments using both natural speech and isolated words as stimuli can be helpful in testing these hypotheses. In addition, lexical competition and segment prediction error are also more correlated in continuous speech. Hence, in order to clearly distinguish the competitive- and predictive-selection accounts, it is necessary to intentionally manipulate lexical probability using a specifically designed paradigm. I will report a MEG study testing the two accounts using the competitor priming paradigm in Chapter 2.

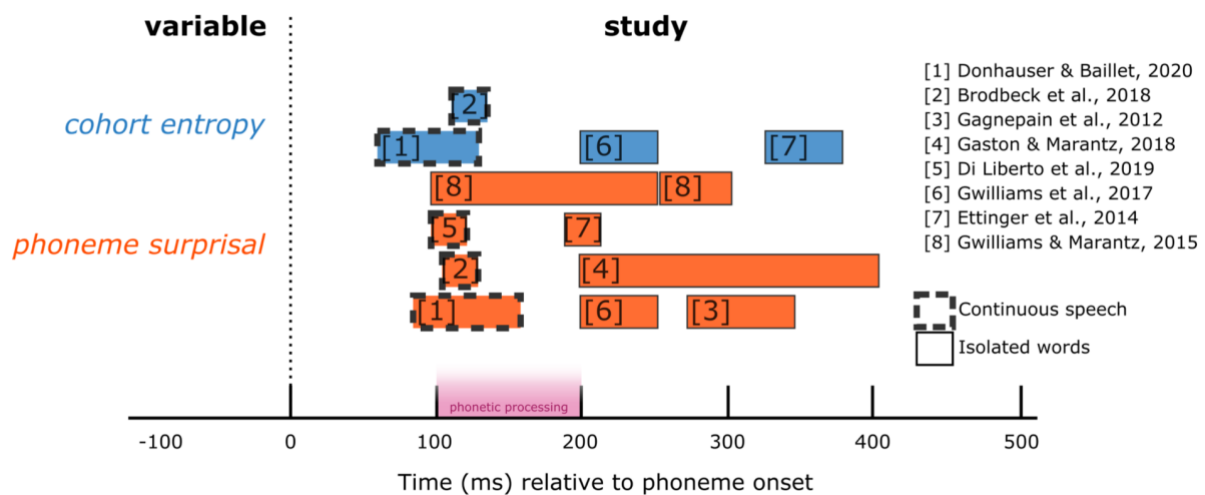


Figure 1.8. A summary of the time course at which different studies of spoken word perception, using either isolated words or continuous speech, found significant effects of cohort entropy (which quantifies competition) and phoneme surprisal (similar to prediction error), figure adapted from Gwilliams & Davis (2021).

1.4 Mechanisms of word memory encoding

In everyday life, human can not only perceive and comprehend speech easily, but also recall the words they heard, even if they are novel words such as *noob*, *troll*, from conversations without intentional effort to remember them. In this section, I will review cognitive theories that explain how subsequent recognition memory can be enhanced and their supporting neural mechanisms. I will first review the effect of prediction on memory, which can be manipulated through the competitor priming paradigm. Then I will focus on the effect of levels of processing and divided attentional resources on memory, which can be modulated by experimental tasks. Finally, I will review the complementary learning systems account which explains how novel information such as pseudowords can be encoded in memory.

1.4.1 Prediction and memory

As reviewed in previous sections, the competitor priming effect during spoken word recognition can be explained by the Predictive Coding account (Davis & Sohoglu, 2020), i.e. competitor-primed words evoke larger prediction error, which cause delay in identifying these words. At the same time, the same prediction error may also have long-lasting effects on their subsequent episodic memory. For example, PIMMS (Predictive Interactive Multiple Memory Systems; Henson & Gagnepain, 2010, see Figure 1.9) proposes that stronger prediction error drives episodic learning due to greater synaptic exchanges between the medial temporal lobe (MTL) and neocortex, hence unexpected items should be better remembered. However, literature from schema-based learning showed that better memory is associated with schema-congruent information that matches representations in semantic memory (Alba & Hasher, 1983; Cycowicz et al., 2008; Hölzje et al., 2019). Conversely, information that is not relevant to the schema is remembered less well (Sweegers et al., 2015), hence unexpectancy may also hinder

memory. Here, I review empirical evidence for both sides of argument, with a focus on word memory studies.

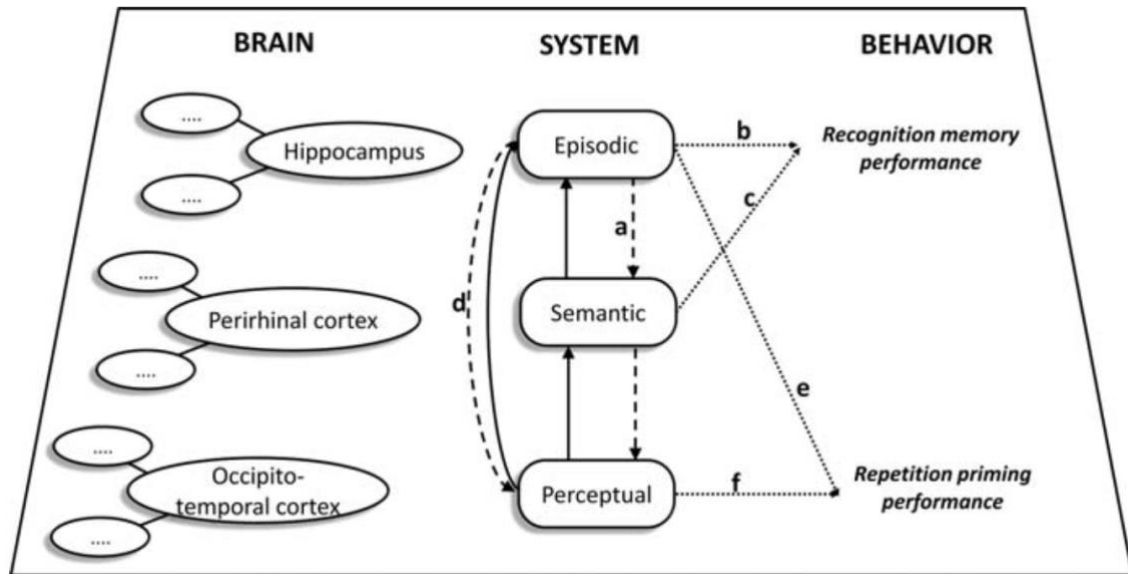


Figure 1.9. Illustration of the relationship between brain, hypothetical memory system and behaviour as proposed by the PIMMS model (from Henson & Gagnepain, 2010). Letters mark different routes that represent possible causes of behavioural outcomes in recognition memory task and perceptual priming paradigm. The outcome of recognition memory task could be affected by interactions between episodic memory system in the hippocampus and semantic memory system in the perirhinal cortex, while perceptual priming task is more related to the perceptual system in the occipito-temporal cortex with potential input from the hippocampal episodic system.

According to PIMMS, memory is organized in a processing hierarchy with multiple levels, in which higher levels predict activities in lower sensory levels. The difference between predictions and lower-level activities, namely the prediction error, updates learning. In Bayesian terms, the higher-level predictions are equivalent to prior probability distribution,

while the lower-level activities are the likelihood (evidence). However, depending on the strength and accuracy of the prediction and whether the sensory input has pre-existing representations (e.g. familiar words or novel pseudowords), the size of prediction error is also different.

For example, PIMMS suggests that there should be a large prediction error when the prediction is based on the initial segments of a familiar word that has been previously presented and that the input is an unexpected competitor target word. This prediction error should lead to better memory of the target word. However, novel items such as pseudowords have no prior phonological and semantic representations, hence even though they evoke large prediction error in perceptual systems (in the context of spoken word recognition), there was only minimal prediction error in the semantic memory system, whose interaction with the episodic system helps driving memory encoding and recollection. Hence, PIMMS does not predict strong memory effects for pseudowords. It may be that many such learning episodes and long-term consolidations are needed before novel items can be effectively remembered (see section 1.4.3 on complementary learning systems).

A number of previous studies have demonstrated that larger prediction error evoked either from form prediction or context prediction helps improve memory for words. Tulving and Kroll (1995) presented a list of random words to participants in the familiarization phase, and then presented half of the words that had appeared previously in the critical study phase, while the other half had not been presented before. They found that participants' memory for the previously unprimed words were better than the primed words. This result can be explained by the PIMMS account as participants might have been expecting to see repeated words during the critical study phase, and that the appearance of unprimed words was more surprising, which generated larger prediction error that facilitated encoding. Indeed, Kafkas and Montaldi (2015a) showed that the proportion of primed and unprimed items in the critical

phase was essential and that the more unexpected type of item (with smaller proportion during the critical phase) was better remembered. In addition, Corley and colleagues (2007) found in their ERP study that words were better remembered when they were shown as an unexpected ending of a sentence compared to when they fit into a predictable sentence (e.g. “That drink’s too hot; I have just burnt my *nails*.” vs “Everyone’s got bad habits and mine is biting my *nails*.”). Unpredictable words also led to greater negativity over 300-500ms (i.e. N400) after word onset mainly over central and midline locations. Furthermore, Haeuser & Kray (2021) used a self-paced reading task which manipulated the predictability of gender-marked German nouns based on prior sentence context, with the pre-nominal gender-marked article (i.e. *the*, *dem* or *der* in German) and buffer adjectives (i.e. *old but reliable*, *alten aber zuverlässig* in German) used as an early cue. They found that early prediction error during the pre-nominal period enhanced subsequent memory for nouns.

In contrast, other studies also showed that more predictable words are better remembered than unexpected ones. Hölte and colleagues (2019) conducted an ERP study in which participants learned category exemplar words that were either congruent (e.g. *dog*) or incongruent (e.g. *pepper*) with a preceding category cue (e.g. a four-footed animal). They found that congruent words were better remembered than incongruent words and that the subsequent memory effects for congruent words emerged in the N400 time window in the frontal and parietal area. Riggs and colleagues (1993) also found that better word subsequent memory tested through free recall, cued recall and multiple-choice recognition was related with more predictable prose passages. These results contradict with the PIMMS account and indicate that information integrating with prior knowledge (i.e. schema) better can be encoded more easily.

However, Greve and colleagues (2019) found evidence for both accounts in which memory was a “U-shaped” function of congruency, with superior memory for both congruent and incongruent trials (at the two sides of “U”) compared to controls (in the middle of “U”), as

supported by the SLIMM framework (van Kesteren et al., 2012). This framework suggests that memory performance can be a nonlinear function of congruency, by which better schematic memory for congruent items is mediated by medial prefrontal cortex (mPFC) and better instance memory for incongruent items is mediated by the MTL. Therefore, it seems that prediction plays a complex role on memory and that a one-sided view cannot cover the full story. In Chapter 4, I report subsequent word memory results after encoding during the word perception phase (see Chapter 2 and 3). With the manipulation of initial sound segment-based predictions rather than schema-based semantic predictions during encoding, the competitor priming provides a different method for testing the PIMMS account.

1.4.2 Levels of processing and attentional resources

Apart from prediction, other mechanisms such as levels of processing (Craik & Lockhart, 1972) and attentional resources (Craik et al., 1996; Jurica & Shimamura, 1999) also affect episodic memory encoding.

According to Craik and Lockhart (1972), words that are processed at the semantic level (the deeper level) are better remembered than words processed at the shallower phoneme or feature level because deep levels of analysis are associated with stronger memory trace. Wagner and colleagues (1998b) conducted an fMRI study in which participants performed a semantic processing task (deciding if a word is abstract or concrete) and a non-semantic task (deciding if a word is printed in upper- or lowercase letters). They found that 85% of words were remembered in the subsequent memory test following the semantic task, while only 47% of words were remembered following non-semantic processing. Both semantic processing (vs non-semantic processing) and remembered (vs forgotten) words evoked greater activation in the left prefrontal cortex, left parahippocampal and fusiform gyri. Similar results were also found by Kapur et al. (1994), who reported that deep encoding task involving semantic

processing (noun categorisation as living or non-living) promoted better subsequent memory than shallow encoding task (detecting the presence of the letter a), and that the positron emission tomography (PET) demonstrated robust left prefrontal activation for deep processing.

Another reason why deeper, more semantic encoding might lead to better subsequent memory is that completing these tasks might require additional attentional resources (e.g. because they are more difficult or attention demanding). According to Craik et al. (1996), an item is more likely to be forgotten when attention was directed away from it. Using PET, Fletcher and colleagues (1995) further showed that adding a distracting task during intentional encoding impaired memory performance and diminished neural activity in the frontal cortex. In addition, Jurica and Shimamura (1999) found that when participants spent more attentional resources in doing the task (i.e. answering questions about topics, rather than just reading statements about topics), their memory for items (i.e. the question topics) was enhanced, although at the same time, their memory for the source of the items (i.e. the person who asked the question) was disrupted. Hence, they proposed that the encoding of item and source information are dissociable and there was an item-source trade-off during episodic encoding.

Overall, it seems that deeper level of processing and greater attentional resources facilitate subsequent memory of items. These processes also seem closely related and are both engaged by frontal activities. However, it is also possible that episodic encoding involves separate processes for item and source memory, and that more attention spent on items does not necessarily enhance memory for forming item-source associations. I will discuss these more in Chapter 4, in which I report memory results that are relevant to manipulations of levels of processing and attentional effort due to the use of lexical and sub-lexical tasks during encoding.

1.4.3 The complementary learning systems for novel word encoding

While the PIMMS account (as reviewed in 1.4.1) does not predict strong encoding of novel items without prior semantic presentations (e.g. pseudowords), it is undeniable that we must be able to learn something about these new items when encountering them, as learning novel words is a common practice in our daily life. This section reviews the complementary learning systems account and its supporting empirical evidence on how novel items such as pseudowords can be encoded in memory.

The complementary learning systems framework (CLS; Kumaran et al., 2016; McClelland et al., 1995) suggests that effective learning requires two complementary systems: one is rapid learning of episodes of individual items and experiences located in the hippocampus, the other is gradual acquisition of structured knowledge about the environment in the neocortex through active rehearsal and inactive states such as sleep (see Figure 1.10). The CLS account also suggests that with further learning in the neocortex, the dependence of the memory on the hippocampus is gradually reduced via the process of systems consolidation. Indeed, such multiple memory systems are demonstrated by amnesic patients with lesions in the hippocampus (e.g. H.M.), who have trouble remembering previous experiences, but have no problem in neocortical memory functions such as motor learning and perceptual learning (Squire, 2009). Crucially, the memory impairment in hippocampal amnesic patients is less severe as the time between memory establishment (prior to the brain damage) and amnesia onset increases (Kapur & Brooks, 1999), which suggests that the knowledge is less dependent on intact hippocampal structure and connectivity.

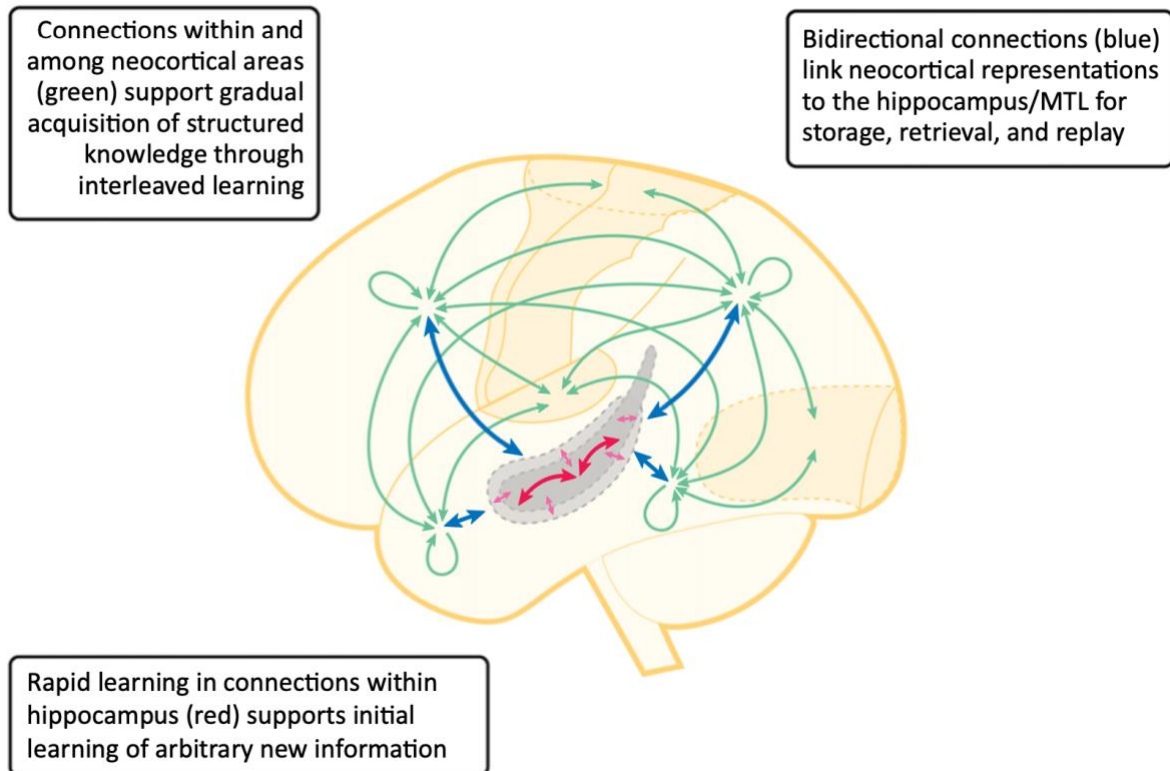


Figure 1.10. A depiction of the complementary learning systems (CLS) represented in the brain, showing connections within the hippocampus for initial rapid learning (red), connections within neocortical areas for gradual acquisition of structured knowledge (green) and interactions between the two systems (blue) (adapted from Kumaran et al., 2016).

Davis and Gaskell (2009) applied this CLS account to spoken word learning. They propose that novel spoken word learning also involves the two systems consistent with CLS. An initial rapid acquisition of novel words happens in the medial temporal lobe (MTL) and, after the initial learning stage, these novel words are consolidated offline slowly in the neocortex, with sleep playing an important role during this process. In addition, in order to allow for recognition of both familiar and novel words, the hippocampal and neocortical routes should operate with prioritization such that the neocortical route is dominant until the point when it fails to recognize familiar words, and this is the point when the hippocampal route

would activate to support rapid learning for novel words. This process is potentially triggered by high prediction error as suggested by Davis & Sohoglu (2020). Hence, such operations of the two routes also mean that novel words will not affect recognition of familiar words until they are consolidated in the neocortex (see also Dumay & Gaskell, 2012).

Neural evidence has shown that hippocampal activation contributes to initial acquisition of novel words both on associative learning of meaning and form-only phonological learning. Breitenstein and colleagues (2005) used fMRI to record participants neural activity during spoken pseudoword-picture associative learning. They found that the strength of the neural activity in hippocampus in the first learning block was already predictive of subsequent learning success of pseudowords. Additionally, linear decrease of neural activity in the left hippocampus was shown over five learning blocks of the same stimuli pairing and that less decline of the hippocampal activity is associated with more efficient learning of the new words. Mestres-Misse and colleagues (2008) also showed in their fMRI study that successful encoding of written pseudowords presented at the end of constrained sentence contexts is associated with neural activity in the parahippocampal gyrus.

Moreover, Davis and colleagues (2009) tested pseudoword form learning using fMRI. They used three sets of spoken words and pseudowords as stimuli, one set was familiarized through phoneme monitoring one day before testing in the scanner, another set was familiarized about 4 hours before scanning, the last set was presented in the scanner only. The testing in the scanner was conducted using the pause detection task. They found that the hippocampus showed significantly stronger activity for untrained pseudowords compared to recently trained ones during the first scanning run, while this difference was non-significant in subsequent scanning sessions. Furthermore, Takashima et al. (2014) also used a phoneme monitoring task to train participants' learning of two sets of spoken pseudowords: one set presented alone with only the phonological information, and one set presented together with pictures of unfamiliar

objects. Recognition memory tests were then conducted both on the training day and the day after while participants' neural activity was recorded using the fMRI. They found greater activation in the medial temporal lobe for both types of pseudowords compared to untrained ones on the first day, while a connectivity analysis indicated that successfully remembered picture-associated pseudowords connected more to the right hippocampus than form-only pseudowords. Overall, these results indicated that the hippocampus plays an important role in the initial encoding of pseudowords, although the degree might be different for form-only learning and meaning-associated learning.

In addition, evidence was also found in support of the role of overnight consolidation on novel word learning. Gaskell and Dumay (2003) examined the effect of lexicalized novel words on recognition of existing words. They taught participants pseudowords sharing initial segments with existing words (eg. *cathedruke* and *cathedral*) through repeated presentations (36 times each novel word) using a phoneme monitoring task. After the encoding phase, the participants were tested on direct recognition of the novel items using a 2-AFC test, which showed excellent effects of learning. The lexicalization of these novel items was also tested both immediately after the encoding phase and a week later using a pause detection task, which reflects the overall level of lexical activity before the location of the pause (Mattys & Clark, 2002). No evidence of increased competition was shown for existing words immediately after the encoding session, but the competition effect became reliable a week later. The dissociation of immediate form learning and delayed lexicalization of novel words is in line with the dual-stage learning process as supported by the CLS account. To further examine the role of sleep on memory consolidation, Dumay and Gaskell (2007) tested two groups of participants, with one group trained at 8am and another group trained at 8pm for novel words learning using the same phoneme monitoring task as in the 2003 study. They tested both groups of participants using 2-AFC, pause detection and free recall immediately after training, 12 hours later and 24

hours later respectively. Both groups showed good recognition memory of the novel items in the 2-AFC test at all time points, however, only after overnight consolidation (i.e. after 24 hours for the 8am training group, after 12 and 24 hours for the 8pm training group) did the lexical competition effects appear in the pause detection task. Similarly, free recall rate was improved only after nocturnal sleep. In addition, other work also showed that similar mechanisms of overnight consolidation operate in children's novel word learning (see James et al., 2017, for review) and that overnight lexical integration is associated with sleep spindle activity (Tamminen et al., 2010).

However, a few studies also suggest that consolidation may not always be required to lexicalize novel words. Fernandes and colleagues (2009) used continuous artificial language as the training stimuli, in which transitional probabilities between adjacent syllables served as the cue to word boundaries. Using the lexical decision task at test, they found effects of lexical interference from familiarized novel words immediately after training. In addition, Kapnoula et al. (2015) used the phoneme monitoring task to familiarize participants with novel words. Immediately after training, they found that splicing a trained novel words (e.g. *nep*) onto the target word (e.g. *net*) caused significantly fewer looks to the target word during eye tracking than splicing an untrained novel word. These studies showed that there are still possibilities of lexicalizing newly learned words without overnight consolidation and the effects may be related to the sensitivity of the tasks used for testing and whether words are trained in isolation (see McMurray et al., 2016, for review).

Overall, most evidence have shown that novel word encoding is associated with rapid initial acquisition in the hippocampus and slow consolidation in the neocortex as suggested by the CLS account. However, since consolidation-related changes are graded, more remains to be learned about the hippocampal plasticity and how different tasks affect novel word learning. I report memory tests in Chapter 4 that investigated the effect of competitor priming on word

and pseudoword memory. Although the role of consolidation on pseudoword memory was not studied directly, we explored how priming of related prior knowledge changes learning and examined the neural underpinnings of the initial encoding of pseudowords.

1.5 Thesis structure

This thesis aims to examine the neural and functional mechanisms of spoken word recognition and memory encoding by using the competitor priming paradigm.

Chapter 2 reports an MEG study that examined the neural mechanisms of spoken word recognition by testing two distinct implementations of Bayes perceptual inference - the competitive-selection accounts (e.g. TRACE) and the predictive-selection accounts (e.g. Predictive-Coding). The hypotheses of these two accounts were tested in terms of the location, timing and direction of the neural responses manipulated through competitor priming.

Chapter 3 describes a series of studies that investigated the effect of lexical and sub-lexical tasks on competitor priming. We explore this by comparing the effect of word/pseudoword lexical decision and pause detection (sub-lexical decisions) during perception of priming and target items. While lexical inhibitory effect and phonological facilitatory effect for similar words are potentially embodied in most spoken word models with hierarchical structures, e.g. TRACE, they are most directly reflected in the Distributed Cohort Model (DCM) with distinct phonological and semantic representations.

Chapter 4 reports tasks and analyses that examined the mechanisms of spoken word memory encoding during competitor priming using additional recognition memory data from the same experiments as reported in Chapter 2 and 3. The competitor priming paradigm changes the degree to which heard words are predicted or not, and hence assesses whether prediction strength and accuracy modulates memory. The use of lexical and sub-lexical

encoding tasks modulated the levels of processing and participants' attention. In addition, the neural responses from the MEG study (described in Chapter 2) that predicted subsequent memory were also analysed to further investigate the neural representations of remembered and forgotten words and pseudowords.

Finally, Chapter 5 summarises and integrates findings reported in previous chapters, with an emphasis on how they advanced our understanding of spoken word recognition and learning. Limitations of these studies and future directions will also be discussed.

2 PREDICTIVE NEURAL COMPUTATIONS SUPPORT

SPOKEN WORD RECOGNITION: EVIDENCE FROM MEG AND COMPETITOR PRIMING

2.1 Introduction

Although it is recognised that humans achieve quick and effortless word recognition by combining current speech input with its prior probability using Bayes theorem (Norris & McQueen, 2008; Davis & Scharenborg, 2016; Gwilliams & Davis, 2021), the underlying neural implementation of Bayesian perceptual inference remains unclear (Aitchison & Lengeyl, 2017). In this chapter, I describe an MEG study that tested two computational accounts of spoken word recognition that both implement Bayes rules.

In competitive-selection accounts (e.g. TRACE, McClelland & Elman, 1986, Figure 2.1A), word recognition is achieved through within-layer lateral inhibition between neural units representing similar words. By this view, *hijack* and *hygiene* compete for identification such that an increase in probability for one word inhibits units representing other similar-sounding words. Conversely, predictive-selection accounts (e.g. Predictive-Coding, Davis & Sohoglu, 2020) suggest that word recognition is achieved through computations of prediction error (Figure 2.1D). On hearing transiently ambiguous speech like /haidʒ/, higher-level units representing matching words make contrasting predictions (/æ/ for *hijack*, /i:n/ for *hygiene*). Prediction error (the difference between sounds predicted and actually heard) provides a signal to update word probabilities such that the correct word can be selected.

In this study, we used the competitor priming effect (Monsell & Hirsh, 1998; Marsolek, 2008), which is directly explicable in Bayesian terms, as competitor priming can increase the

prior probability of the prime competitor word due to the observation of a new instance of the prime word and this slows down identification of the target word. The assumption here is that the change of prior probability caused by competitor priming is achieved through the same mechanism that updates word frequency knowledge. Monsell and Hirsh (1998) showed that this effect can last for 1-5 minutes, such that the recognition of the target word (*hygiene*) is delayed if a competitor word (*hijack*) was presented up to 5 minutes earlier. Therefore, the study reported in this chapter also adopted a prime lag of similar length. However, it is also true that this effect may decay if the lag between prime and target words is too long. For example, Rodd and colleagues (2016) found that word-meaning priming for the selected meanings of ambiguous words was reduced to zero after about 4 hours of delay. Competitor priming effects with a longer lag between prime and target items are further explored in Chapter 3.

The delayed recognition caused by competitor priming could be due to increased lateral inhibition according to competitive-selection accounts, or greater prediction error based on predictive-selection accounts. Thus, similar behavioural effects of competitor priming are predicted by two distinct neural computations (Spratling, 2008). To distinguish them, it is critical to investigate neural data that reveals the direction, timing and level of processing at which competitor priming modulates neural responses. Existing neural data remains equivocal with some evidence consistent with competitive-selection (Bozic et al., 2010; Okada & Hickok, 2006), predictive-selection (Gagnepain et al, 2012), or both mechanisms (Brodbeck et al., 2018; Donhauser et al., 2019). We followed these studies in correlating two computational measures with neural activity: lexical entropy (competitive-selection) and segment prediction error (or phoneme surprisal, for predictive-selection).

We used MEG to record the location and timing of neural responses during spoken words recognition in a competitor priming experiment. Pseudowords (e.g. *hijure*) were included in

our analysis to serve as a negative control for competitor priming, since existing research found that pseudowords neither produce nor show this effect (Monsell & Hirsh, 1998). We compared items with the same initial segments (words *hygiene*, *hijack*, pseudowords *hijure*, *higent* share /haidʒ/) and measured neural and behavioural effects concurrently to link these two effects for single trials.

While lexical entropy and prediction error are correlated for natural speech (see Gwilliams & Davis, 2021), this competitor priming manipulation allows us to make differential predictions as illustrated in Figure 2.1. As shown in panel A, B and C, the TRACE model predicts that spoken word recognition is achieved at the word level through lexical competition (as quantified by lexical entropy), while panel D, E and F indicate that, in the Predictive Coding account, it is the segment prediction error calculated at the phoneme level that drives spoken word recognition. Although our hypotheses for the MEG neural responses are based on these assumptions, it should also be noted that both accounts are hierarchical and that the phoneme level of the TRACE model and the lexical level of the Predictive Coding model are both crucial in contributing towards the process of word recognition, hence these assumptions could easily be revised depending on the measurement being used. Indeed, Luthra and colleagues (2021) showed that TRACE simulation based on phoneme activation may also explain signal reduction when input is consistent with expectations just as the Predictive Coding account.

Overall, our predictions are the following: (1) before the deviation point (DP, the point at which similar-sounding words diverge), competitor priming increases lexical entropy and hence neural responses (Figure 2.1B&C Pre-DP). Such responses might be observed in inferior frontal regions (Zhuang et al., 2011) and posterior temporal regions (Prabhakaran et al., 2006). However, prediction error will be reduced for pre-DP segments, since heard segments are shared and hence more strongly predicted (Figure 2.1E&F Pre-DP). This should be reflected in the superior temporal gyrus (STG, Sohoglu & Davis, 2016). (2) After the DP, predictive-

selection but not competitive-selection accounts propose that pseudowords evoke greater signals in the left-STG, since they evoke maximal prediction errors (Figure 2.1E&F Pseudoword, Post-DP). (3) Furthermore, in predictive-selection theories, competitor priming is associated with an increased STG response to post-DP segments due to enhanced prediction error caused by mismatch between primed words (predictions) and heard speech (Figure 2.1E&F Word, Post-DP).

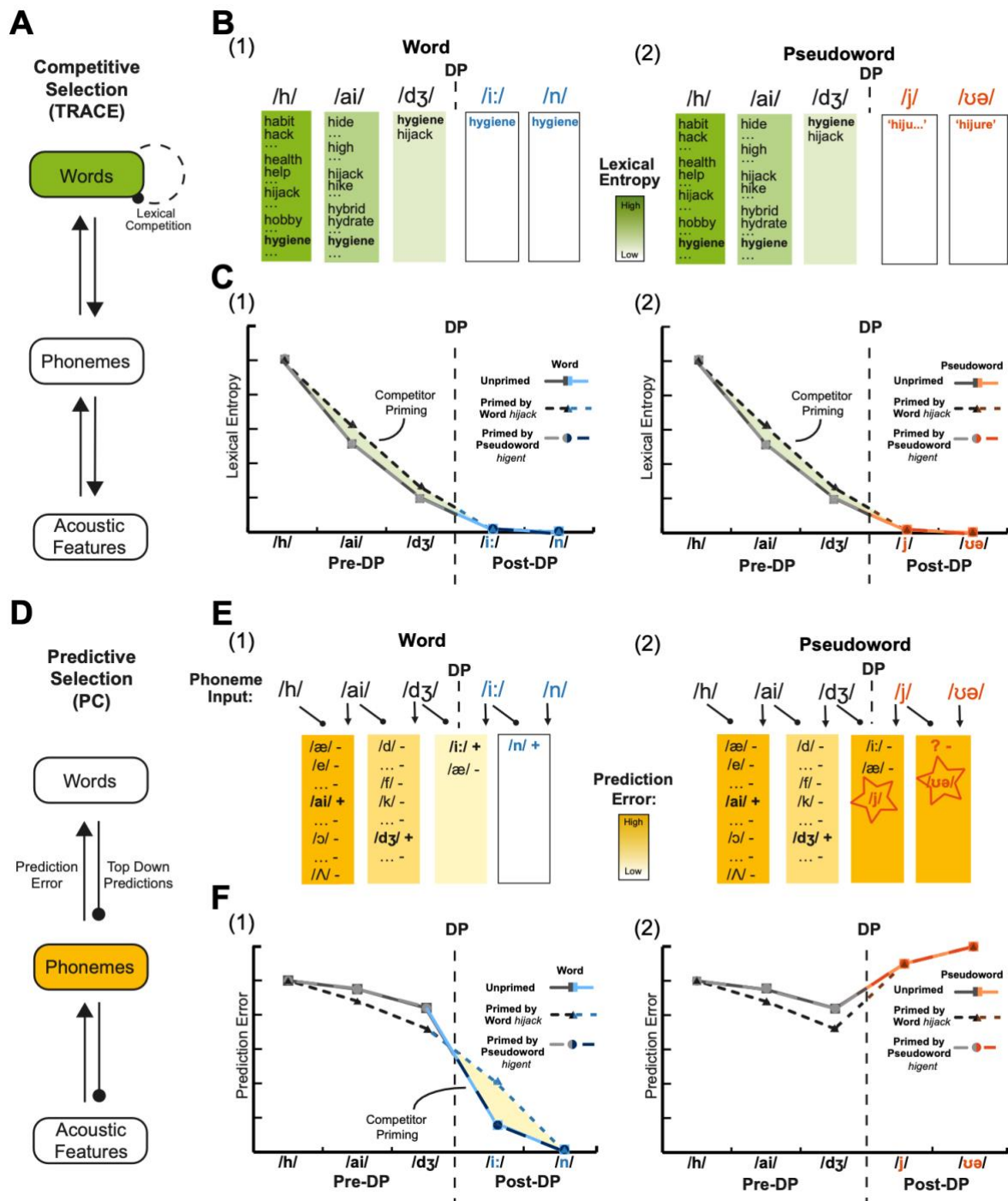


Figure 2.1 Illustration of neural predictions based on competitive-selection and predictive-selection models respectively for recognition of a word (*hygiene*) or pseudoword (*hijure*) that is unprimed or primed by a similar-sounding word (*hijack*) or pseudoword (*higent*). **A.** In a competitive-selection model, such as TRACE (McClelland & Elman, 1986), word recognition is achieved through within-layer lexical competition. **B.** Illustration of the competitive-selection procedure for word (e.g. *hygiene*) and pseudoword (e.g. *hijure*) recognition. Phoneme input triggers the activation of multiple words beginning with the same segments, which compete with each other until one word is selected. No word can be selected when hearing a pseudoword, though it would be expected that lexical probability (although not lexical entropy) should be greater for words than for pseudowords. **C.** Illustration of neural predictions based on lexical entropy. Lexical entropy gradually reduces to zero as more speech is heard. Before the deviation point (hereafter DP) at which the prime (*hijack*) and target (*hygiene*) diverge, these items are indistinguishable, and competitor priming should transiently increase lexical entropy (shaded area). After the DP, competitor priming should not affect entropy since prime and target words can be distinguished. **D.** In a predictive-selection model such as the Predictive-Coding account (PC, Davis & Sohoglu, 2020), words are recognised by minimising prediction error, which is calculated by subtracting the predicted segments from the current sensory input. **E.** Illustration of the predictive-selection procedure during word (e.g. *hygiene*) and pseudoword (e.g. *hijure*) recognition. Speech input evokes predictions for the next segment (based on word knowledge as in panel B), which is then subtracted from the speech input and used to generate prediction errors that update lexical predictions (+ shows confirmed predictions that increase lexical probability, - shows disconfirmed predictions that decrease lexical probability). **F.** Illustration of neural predictions based on segment prediction error. Before the DP, priming of initial word segments should strengthen predictions and reduce prediction error. After the DP, there will be greater mismatch between predictions and heard

speech for competitor-primed words and hence primed words should evoke greater prediction error than unprimed words (shaded area). This increased prediction error should still be less than that observed for pseudowords, which should evoke maximal prediction error regardless of competitor priming due to their post-DP segments being entirely unpredictable.

2.2 Methods

2.2.1 Participants

Twenty-four (17 female, 7 male) right-handed, native English speakers were tested after giving informed consent under a process approved by the Cambridge Psychology Research Ethics Committee. This sample size was selected based on previous studies measuring similar neural effects with the same MEG system (Gagnepain et al. 2012; Sohoglu & Davis, 2016; Sohoglu et al. 2012, etc.). All participants were aged 18-40 years and had no history of neurological disorder or hearing impairment based on self-report. Two participants' MEG data were excluded from subsequent analyses respectively due to technical problems and excessive head movement, resulting in 22 participants in total. All recruited participants received monetary compensation.

2.2.2 Experimental Design

To distinguish competitive- and predictive-selection accounts, we manipulated participants' word recognition process by presenting partially mismatched auditory stimuli prior to targets. Behavioural responses and MEG signals were acquired simultaneously. Prime and target stimuli pairs form a repeated measures design with two factors (lexicality and prime type). The

lexicality factor has 2 levels: word and pseudoword, while the prime type factor contains 3 levels: unprimed, primed by same lexical status, primed by different lexical status. Hence the study is a factorial 2 x 3 design with 6 conditions: unprimed word (*hijack*), word-primed word (*hijack-hygiene*), pseudoword-primed word (*basef-basin*), unprimed pseudoword (*letto*), pseudoword-primed pseudoword (*letto-lettan*), word-primed pseudoword (*boycott-boymid*). Prime-target pairs were formed only by stimuli sharing the same initial segments. Items in the two unprimed conditions served as prime items in other conditions and they were compared with target items (Figure 2.2A).

The experiment used a lexical decision task (Figure 2.2B) implemented in MATLAB through Psychtoolbox-3 (Kleiner et al. 2007), during which participants heard a series of words and pseudowords while making lexicality judgments to each stimulus by pressing buttons using their left index and middle fingers only, with the index finger pressing one button indicating word and the middle finger pressing the other button indicating pseudoword. 344 trials of unique spoken items were presented every ~3 seconds in two blocks of 172 trials, each block lasting approximately 9 minutes. Each prime-target pair was separated by 20 to 76 trials of items that do not start with the same speech sounds, resulting in a relatively long delay of ~1-4 minutes between presentations of phonologically-related items. This delay was chosen based on Monsell and Hirsh (1998), who suggest that it prevents strategic priming effects (Norris et al. 2002). Stimuli from each of the quadruplets were Latin-square counterbalanced across participants, i.e. stimulus quadruplets that appeared in one condition for one participant were allocated to another condition for another participant. The stimulus sequences were pseudo-randomised using Mix software (van Casteren & Davis, 2006), so that the same type of lexical status (word/pseudoword) did not appear successively on more than 4 trials.

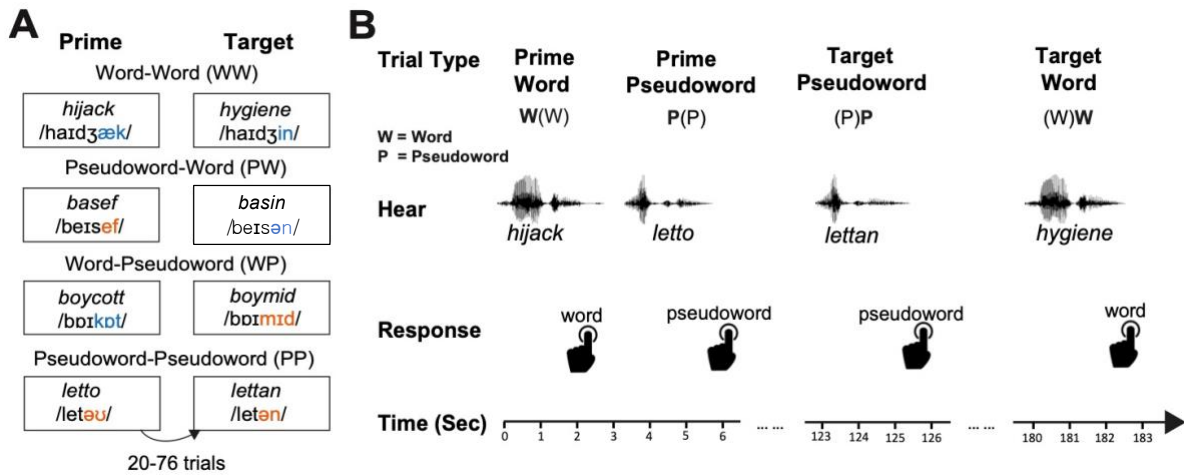


Figure 2.2 Experimental design. **A.** Four different types of prime-target pairs. Each pair was formed by two stimuli from the same quadruplet, separated by between 20 to 76 trials of items that do not share the same initial segments. **B.** Lexical decision task. Participants made lexicality judgments to each item they heard via left hand button-press. The response time was recorded from the onset of the stimuli. As shown, items within each quadruplet are repeated after a delay of ~1-4 minutes following a number of other intervening stimuli.

2.2.3 Stimuli

The stimuli consisted of 160 sets of four English words and pseudowords (see Appendix A), with durations ranging from 372 to 991 ms ($M = 643$, $SD = 106$). Each set contained 2 words (e.g. *letter*, *lettuce*) and 2 phonotactically-legal pseudowords (e.g. *letto*, *lettan*) that share the same initial segments (e.g. /let/) but diverge immediately afterwards.

We used polysyllabic word pairs ($M_{\text{syllable}} = 2.16$, $SD_{\text{syllable}} = 0.36$) instead of monosyllabic ones in our experiments so as to identify a set of optimal lexical competitors that are similar to their prime yet dissimilar from all other items. All words were selected from the CELEX database (Baayen et al., 1993). Their frequencies were taken from SUBTLEX UK corpus (Van

Heuven et al., 2014) and restricted to items under 5.5 based on log frequency per million word (Zipf scale, Van Heuven et al., 2014). In order to ensure that any priming effect was caused purely by phonological but not semantic similarity, we also checked that all prime and target word pairs have a semantic distance of above 0.7 on a scale from 0 to 1 based on the Snaut database of semantic similarity scores (Mandera et al., 2017), such that morphological relatives (e.g. darkly/darkness) were excluded.

All spoken stimuli were recorded onto a Marantz PMD670 digital recorder by a male native speaker of southern British English in a sound-isolated booth at a sampling rate of 44.1 kHz. Special care was taken to ensure that shared segments of stimuli were pronounced identically (any residual acoustic differences were subsequently eliminated using audio morphing as described below).

The point when items within each quadruplet begin to acoustically differ from each other is the deviation point (hereafter DP, see Figure 2.3A). Pre-DP length ranged from 150 to 672 ms ($M = 353$, $SD = 96$), while post-DP length ranged from 42 to 626 ms ($M = 290$, $SD = 111$, see Figure 2.3B). Epochs of MEG data were time-locked to the DP. Using phonetic transcriptions (phonDISC) in CELEX, the location of the DP was decided based on the phoneme segment at which items within each quadruplet set diverge ($M_{\text{seg}}=3.53$, $SD_{\text{seg}}=0.92$). To determine when in the speech files corresponds to the onset of the first post-DP segment, we aligned phonetic transcriptions to corresponding speech files using the WebMAUS forced alignment service (Kisler et al., 2017; Schiel, 1999). In order to ensure that the pre-DP portion of the waveform was acoustically identical, we cross-spliced the pre-DP segments of the 4 stimuli within each quadruplet and conducted audio morphing to combine the syllables using STRAIGHT (Kawahara, 2006) implemented in MATLAB. This method decomposes speech signals into source information and spectral information, and permits high quality speech re-synthesis based on modified versions of these representations. This enables flexible averaging

and interpolation of parameter values that can generate acoustically intermediate speech tokens (see Rogers & Davis, 2017, for example). In the present study, this method enabled us to present speech tokens with entirely ambiguous pre-DP segments, and combine these with post-DP segments without introducing audible discontinuities or other degradation in the speech tokens. This way, phonological co-articulation in natural speech was reduced to the lowest level possible at the DP, hence any cross-stimuli divergence evoked in neural responses can only be caused by post-DP deviation.

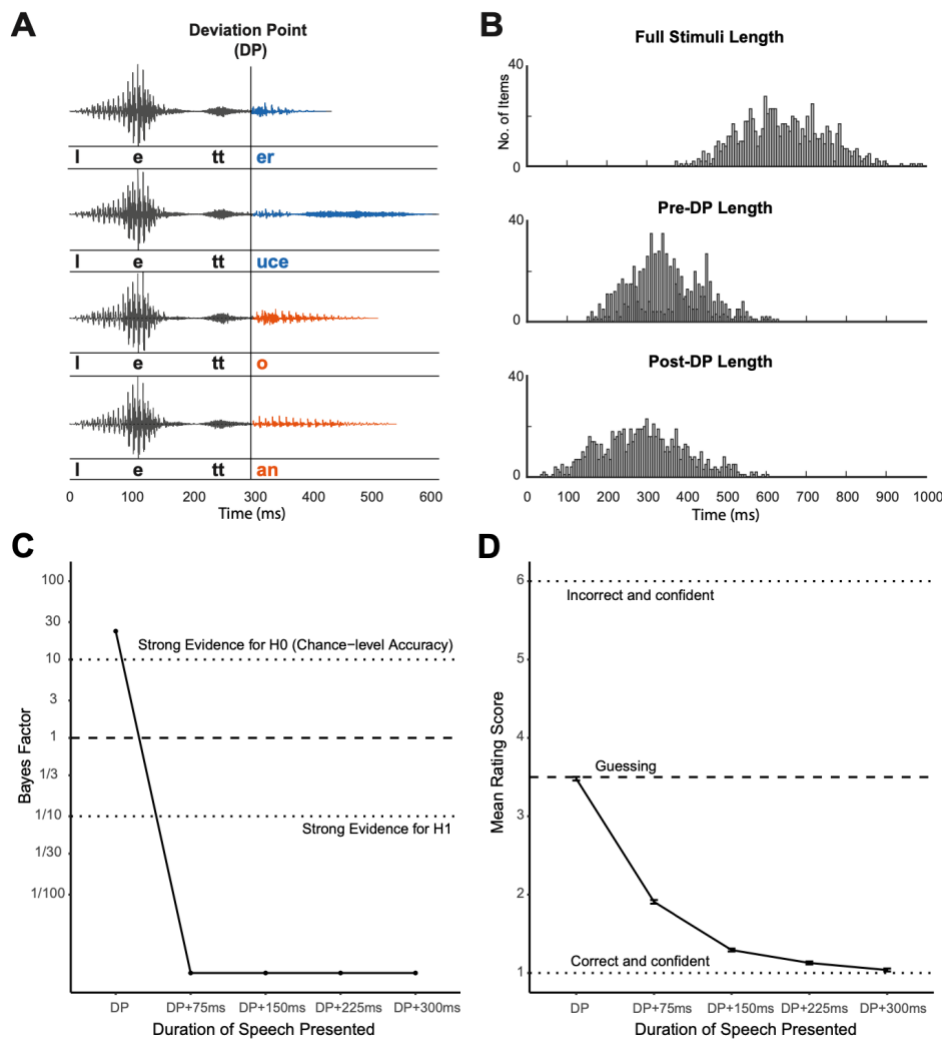


Figure 2.3 Stimuli and post-test gating study results. **A**. Stimuli within the same quadruplet have identical onsets in STRAIGHT parameter space (Kawahara, 2006) and thus only diverge from each other after the deviation point (DP). MEG responses were time-locked to the DP. **B**.

Stimuli length histogram. **C.** Bayes factor for chance level accuracy (BF01) at each post-DP alignment point of the stimuli in the post-test gating study. **D.** Mean rating score at each post-DP alignment point of the stimuli in the gating study.

2.2.4 Post-test Gating Study

As encouraged by a reviewer, we conducted a post-test perceptual experiment using a gating task in order to confirm that the cross-splicing and morphing of our stimuli worked as expected. This experiment used a gating task implemented in JavaScript through JSpsych (de Leeuw, 2015). During the experiment, auditory segments of all 160 pairs of words used in the MEG study were played. Twenty British English speakers were recruited through Prolific Academic online with monetary compensation. The sample size was selected based on a similar gating study conducted by Davis et al. (2002). Participants were evenly divided into two groups, one group were presented with 160 stimuli words with different pre-DP segments (e.g. *hygiene*), while the other group were presented with the other paired 160 stimuli (e.g. *hijack*). Therefore, participants only ever heard one of the two items in each pair. Stimuli segments of each word item consist of the pre-DP segment and, depending on the stimuli length, also longer segments that are 75ms, 150ms, 225ms and 300ms post DP. The segments of each word were presented in a gating manner, with the shortest segment played the first and the full item played at the end. After hearing each segment (e.g. /haidʒ/), participants were also presented with the writing of the word (e.g. *hygiene*) that contained the segment and the other paired word that shared the same pre-DP segment (e.g. *hijack*) on the screen. We asked the participants to choose which item the auditory segment matches and indicate their confidence from a rating scale of 1 to 6, with 1 representing being very confident that the item is the one on the left and 6 representing being very confident that the item is the one on the right, while 3 and 4 representing guessing

the possible item. In order to avoid potential practice effect, we also added 40 filler stimuli that are identifiable on initial presentation.

Given our goal of assessing whether there is any information to distinguish the words prior to the divergence point, we needed to adopt an analysis approach that could confirm the null hypothesis that no difference exists between perception of the shared first syllable of word pairs like *hijack* and *hygiene*. We therefore analysed the results using Bayesian methods which permit this inference. Participants' response accuracy was analysed using mixed-effect logistic regression and confidence rating scores were analysed using mixed-effect linear regression using the *brms* package (Bürkner, 2017) implemented in R. Response scores were computed in a way such that correct and most confident responses were scored 1, while incorrect and most confident responses were scored 6 and so on. Participants and items were included as random factors of the models and there was no fixed factor since we are only interested in the intercepts, whose estimates indicate the logit transformed proportion of correctness in the logistic model and the mean rating in the linear model respectively. We chose weakly informative priors for each model and conducted Bayes Factor analyses through the Savage-Dickey density ratio method (Wagenmakers et al., 2010). Model estimate, standard error, lower and upper boundary of 95% credible interval (CI) are also reported.

When checking our data, we found that 16 pairs of word items were not morphed correctly, hence the spectral information of the pre-DP segments of these word pairs were not exactly the same and some of them diverged acoustically before the DP due to coarticulation. Therefore, we excluded these items from analyses of the gating data and confirmed that excluding these items did not modify the interpretation or significance of the MEG or behavioral results reported in the paper.

As shown in Figure 2.3C, we found that when gating segments ended at the DP, Bayes factor provides strong evidence in favour of the null hypothesis, chance-level accuracy (i.e.

proportion of correct responses is 0.5), $\beta = 0.04$, $SE = 0.08$, $lCI = -0.11$, $uCI = 0.20$, $BF_{01} = 23.04$. This indicates that participants could not predict the full stimuli based on hearing the pre-DP segments. On the other hand, the Bayes factor at later alignment points is close to 0, providing extremely strong evidence for the alternative hypothesis that the proportion of correct responses is higher than 0.5 (75ms post-DP: $\beta = 3.41$, $SE = 0.22$, $lCI = 2.99$, $uCI = 3.85$, $BF_{01} < 0.01$; 150ms post-DP: $\beta = 6.26$, $SE = 0.56$, $lCI = 5.24$, $uCI = 7.41$, $BF_{01} < 0.01$; 225ms post-DP: $\beta = 7.39$, $SE = 1.02$, $lCI = 5.65$, $uCI = 9.72$, $BF_{01} < 0.01$; 300ms post-DP: $\beta = 8.04$, $SE = 1.88$, $lCI = 4.99$, $uCI = 12.32$, $BF_{01} < 0.01$). Figure 2.3D shows that, with the gating segment becoming longer, the rating scores gradually reduce (lower scores indicating more accurate and more confident identification). We examined whether the mean score at the DP is equal to 3.5 (i.e. chance performance) and found strong evidence supporting the null hypothesis, $\beta = -0.02$, $SE = 0.04$, $lCI = -0.10$, $uCI = 0.06$, $BF_{01} = 21.79$, which is consistent with the accuracy results. Furthermore, in order to refine the estimate of the time point at which participants recognise the stimuli with enough confidence, we also investigated at what alignment point is there evidence showing the mean score lower than 2 (i.e. participants indicating more confident identification). We found moderate evidence supporting the null hypothesis (mean score equals to 2) at 75ms post-DP ($\beta = -0.09$, $SE = 0.08$, $lCI = -0.25$, $uCI = 0.07$, $BF_{01} = 6.07$), but extremely strong evidence in favour of the alternative hypothesis at 150ms post-DP ($\beta = -0.71$, $SE = 0.05$, $lCI = -0.79$, $uCI = 0.62$, $BF_{01} < 0.01$). These results show that critical acoustic information that supports confident word recognition arrives between 75ms and 150ms post-DP.

Overall, the post-test gating study confirmed that the pre-DP segments of correctly morphed stimuli are not distinguishable within each stimuli set. However, since we found items that were not correctly morphed during this control study, we did a thorough check of our stimuli and identified all the problematic items (16 words and 12 pseudowords), which resulted

in 8.68% of all trials presented in the MEG study. In order to double check our MEG study results, we then removed all these problematic trials from the data and reanalysed the data using the same methods as described in the method section. Fortunately, we did not find any inconsistent pattern or significance in our behavioural or neural results compared to those reported with all trials included (see Appendix B). Therefore, I report the original MEG and behavioural results with all items included in this Chapter.

2.2.5 Behavioural Data Analyses

Response times (RTs) were measured from the onset of the stimuli and inverse-transformed so as to maximise the normality of the data and residuals; Figures report untransformed response times for clarity. Inverse-transformed RTs and error rates were analysed using linear and logistic mixed-effect models respectively using the lme4 package in R (Bates et al. 2014). Lexicality (word, pseudoword) and prime type (unprimed, primed by same lexical status, primed by different lexical status) were fixed factors, while participant and item were random factors. Maximal models accounting for all random effects were attempted wherever possible, but reduced random effects structures were applied when the full model did not converge (Barr et al., 2013). Likelihood-ratio tests comparing the full model to a nested reduced model using the Chi-Square distribution were conducted to evaluate main effects and interactions. Significance of individual model coefficients were obtained using t (reported by linear mixed-effect models) or z (reported by logistic mixed-effect models) statistics in the model summary. One-tailed t statistics for RTs are also reported for two planned contrasts: (1) word-primed versus unprimed conditions for word targets, and (2) word-primed versus pseudoword-primed conditions for word targets.

When assessing priming effects, we excluded data from target trials in which the participant made an error in the corresponding prime trial, because it is unclear whether such

target items will be affected by priming given that the prime word was not correctly identified. In addition, three trials with RTs shorter than the average pre-DP length (353ms) were removed from further analysis, since responses before words and pseudowords acoustically diverge are too quick to be valid lexical decision responses.

2.2.6 MEG Data Acquisition, Processing and Analyses

Magnetic fields were recorded with a VectorView system (Elekta Neuromag) which contains a magnetometer and two orthogonal planar gradiometers at each of 102 locations within a hemispherical array around the head. Although electric potentials were recorded simultaneously using 68 Ag-AgCl electrodes according to the extended 10-10% system, these EEG data were excluded from further analysis due to excessive noise. All data were digitally sampled at 1 kHz. Head position were monitored continuously using five head-position indicator (HPI) coils attached to the scalp. Vertical and horizontal electro-oculograms were also recorded by bipolar electrodes. A 3D digitizer (FASTRAK; Polhemus, Inc.) was used to record the positions of three anatomical fiducial points (the nasion, left and right preauricular points), HPI coils and evenly distributed head points for use in source reconstruction.

MEG Data were preprocessed using the temporal extension of Signal Source Separation in MaxFilter software (Elekta Neuromag) to reduce noise sources, normalise the head position over blocks and participants to the sensor array and reconstruct data from bad MEG sensors. Subsequent processing was conducted in SPM12 (<https://www.fil.ion.ucl.ac.uk/spm/>) and FieldTrip (<http://www.fieldtriptoolbox.org/>) software implemented in MATLAB. The data were epoched from -1100 to 2000ms time-locked to the DP and baseline corrected relative to the -1100 to -700ms prior to the DP, which is a period before the onset of speech for all stimuli (Figure 2.3B). Low-pass filtering to 40 Hz was conducted both before and after robust averaging across trials (Litvak et al., 2011). A time window of -150 to 0ms was defined for

pre-DP comparisons based on the shortest pre-DP stimuli length. A broad window of 0 to 1000ms was defined for post-DP comparisons, which covered the possible period for lexicality and prime effects. After averaging over trials, an extra step was taken to combine the gradiometer data from each planar sensor pair by taking the root-mean square (RMS) of the two amplitudes.

Sensor data from magnetometers and gradiometers were analysed separately. We converted the sensor data into 3D images (2D sensor x time) and performed F tests for main effects across sensors and time (the term “sensors” denotes interpolated sensor locations in 2D image space). Reported effects were obtained with a cluster-defining threshold of $p < .001$, and significant clusters identified as those whose extent (across space and time) survived $p < 0.05$ FWE-correction using Random Field Theory (Kilner & Friston, 2010). Region of interest (ROI) analyses for the priming effect were then conducted over sensors and time windows that encompassed the significant pseudoword>word cluster, orthogonal to priming effects. When plotting waveforms and topographies, data are shown for sensors nearest to the critical points in 2D image space.

Apart from the two planned contrasts mentioned above (see Behavioural Data Analyses), which were applied to post-DP analysis, one-tailed t statistics was also reported on the pre-DP planned contrast between unprimed and word-primed items.

2.2.7 Source Reconstruction

In order to determine the underlying brain sources underlying the sensor-space effects, source reconstruction was conducted using SPM’s Parametric Empirical Bayes framework (Henson et al., 2011). To begin with, we obtained T1-weighted structural MRI (sMRI) scans from each participant on a 3T Prisma system (Siemens, Erlangen, Germany) using an MPRAGE sequence.

The scan images were segmented and normalised to an MNI template brain in MNI space. The inverse of this spatial transformation was then used to warp canonical meshes derived from that template brain back to each subject's MRI space (Mattout et al., 2007). Through this procedure, canonical cortical meshes containing 8196 vertices were generated for the scalp and skull surfaces. We coregistered the MEG sensor data into the sMRI space for each participant by using their respective fiducials, sensor positions and head-shape points (with nose points removed due to the absence of the nose on the T1-weighted MRI). Using the single shell model, the lead field matrix for each sensor was computed for a dipole at each canonical cortical mesh vertex, oriented normal to the local curvature of the mesh.

Source inversion was performed with all conditions pooled together using the 'IID' solution, equivalent to classical minimum norm, fusing the magnetometer and gradiometer data (Henson et al, 2011). The resulting inversion was then projected onto wavelets spanning frequencies from 1 to 40 Hz and from -150 to 0ms time samples for pre-DP analysis and 400 to 900ms for post-DP analysis. This post-DP time window was defined by overlapping temporal extent of the pseudoword > word cluster between gradiometers and magnetometers. The total energy within these time-frequency windows was summarised by taking the sum of squared amplitudes, which was then written to 3D images in MNI space.

Reported effects for source analyses were obtained with a cluster-defining threshold of $p < 0.05$ (FWE-corrected). And as in sensor space, ROI analyses were conducted over significant sensors and time windows from the orthogonal pseudoword>word cluster. Factorial ANOVA were carried out on main effects and one-tailed paired t -tests on planned contrasts (see section 2.2.6).

2.3 Results

2.3.1 Behaviour

Response Times. As shown in Figure 2.4A, factorial analysis of lexicality (word, pseudoword) and prime type (unprimed, primed by same lexical status, primed by different lexical status) indicated a significant main effect of lexicality, in which RTs for pseudowords were significantly longer than for words, $X^2(3) = 23.60$, $p < .001$. In addition, there was a significant interaction between lexicality and prime type, $X^2(2) = 10.73$, $p = .005$. This interaction was followed up by two separate one-way models for words and pseudowords, which showed a significant effect of prime type for words, $X^2(2) = 10.65$, $p = .005$, but not for pseudowords, $X^2(2) = 1.62$, $p = .445$. Consistent with the competitor priming results from Monsell and Hirsh (1998), words that were primed by another word sharing the same initial segments were recognised significantly more slowly than unprimed words (for mean raw RTs see Figure 2.4A), $\beta = 0.02$, $SE = 0.01$, $t(79.69) = 3.33$, $p < .001$, and more slowly than pseudoword-primed words, $\beta = 0.02$, $SE = 0.01$, $t(729.89) = 2.37$, $p = .009$. As mentioned earlier (see Introduction), both competitive- and predictive-selection models predicted longer response times to word-primed target words compared to unprimed words, it is hence critical to distinguish the two accounts through further investigation of the MEG responses.

Accuracy. Figure 2.4B shows that there was a trend towards more lexical decision errors in response to words than to pseudowords, although this lexicality effect was marginal, $X^2(3) = 7.31$, $p = .063$. The error rates for words and pseudowords were also affected differently by priming, as indicated by a significant interaction between lexicality and prime type, $X^2(2) = 6.08$, $p = .048$. Follow-up analyses using two separate models for each lexicality type showed there was a main effect of prime type for words, $X^2(2) = 13.95$, $p < .001$, but not for pseudowords, $X^2(2) = 1.93$, $p = .381$. Since we had not anticipated these priming effects on

accuracy, post-hoc pairwise z tests were Bonferroni corrected for multiple comparisons. These showed that pseudoword priming reliably increased the error rate compared to the unprimed condition, $\beta = 1.68$, $SE = 0.54$, $z = 3.14$, $p = .005$, and to the word-primed condition, $\beta = 2.74$, $SE = 0.89$, $z = 3.07$, $p = .007$. Although no specific predictions on accuracy were made a priori by either competitive- or predictive-selection model, it is worth noting that participants might have expected pseudowords to be repeated given the increased error rate of responses to pseudoword-primed target words.

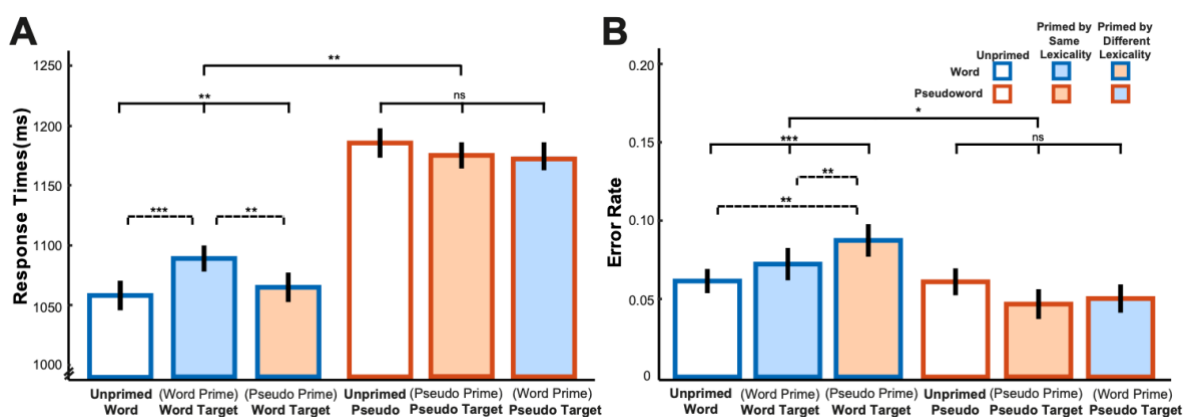


Figure 2.4 Response time results (**A**) and accuracy results (**B**) of the lexical decision task. Bars are color-coded by lexicality and prime type on the x axis (words, blue frame; pseudowords, orange frame; unprimed, no fill; primed by same lexicality, consistent fill and frame colors; primed by different lexicality, inconsistent fill and frame colors). Bars show the subject grand averages, error bars represent \pm within-subject SE, adjusted to remove between-subjects variance (Cousineau, 2005). Statistical significance is shown based on (generalised) linear mixed-effects regression: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Statistical comparisons shown with solid lines indicate the lexicality by prime-type interaction and main effects of prime-type for each lexicality, whereas comparisons with broken lines indicate the significance of pairwise comparisons.

2.3.2 MEG

In order to explore the impact of lexicality and competitor priming on neural responses to critical portions of speech stimuli, both before and after they diverge from each other, MEG responses were time-locked to the DP. All reported effects are family-wise error (FWE)-corrected at cluster level for multiple comparisons across scalp locations and time at a threshold of $p < 0.05$. We reported data from gradiometers, magnetometers and source space wherever possible, since sensor x time analyses help define the time-windows used by source localisation. Although some minor effects were shown in only one of these analyses, our most interesting effects are reliable in all three data types.

Pre-DP analyses. We assessed neural responses before the DP, during which only the shared speech segments have been heard and hence the words and pseudowords in each stimulus set are indistinguishable. Since there could not have been any effect of lexical status pre-DP, only prime type effects were considered in this analysis. Predictive- and competitive-selection accounts make opposite predictions for pre-DP neural signals evoked by word-primed items compared to unprimed items. We therefore conducted an F-test for neural differences between these two conditions across the scalp and source spaces over a time period of -150 to 0ms before the DP. A significant cluster of 295 sensor x time points ($p = .023$) was found in gradiometers over the mid-left scalp locations from -28 to -4ms (Figure 2.5A), in which unprimed items evoked significantly greater neural responses than word-primed items. On the suggestion of a reviewer, and mindful of the potential for these pre-DP neural responses to be modulated by post-DP information, we report an additional analysis with a lengthened analysis time window of -150ms to 100ms. Again, we found a significant unprimed > word-primed cluster of 313 sensor x time points ($p = .033$) over the exact same locations in gradiometers from -28 to -3ms pre-DP, which confirmed that this pre-DP effect was not pushed forward by any post-DP effect. We did not find any cluster showing stronger neural responses for word-

primed items than unprimed items and no clusters survived correction for multiple comparisons for magnetometer responses or for analysis in source space.

To further examine these results, we also conducted ROI analysis of gradiometer signals evoked by unprimed and primed items averaged over the same -150 to 0ms pre-DP time window but across the scalp locations that showed the post-DP lexicality effect at which pseudowords elicited greater neural responses than words (see Figure 2.6A). As shown in Figure 2.5B, the results indicated that unprimed items elicited significantly stronger neural responses than word-primed items, $t(21) = 2.41$, $p = .013$, consistent with the whole-brain analysis. In particular, the mid-left cluster shown in panel A partially overlaps with the post-DP pseudoword>word cluster. The direction and location of these pre-DP neural responses are in accordance with the predictive-selection account and inconsistent with the competitive-selection account. A surprising finding is that post-hoc analysis also showed greater neural responses evoked by unprimed items than pseudoword-primed items, $t(21) = 2.69$, $p = .014$, although we had not predicted these effects from pseudoword primes.

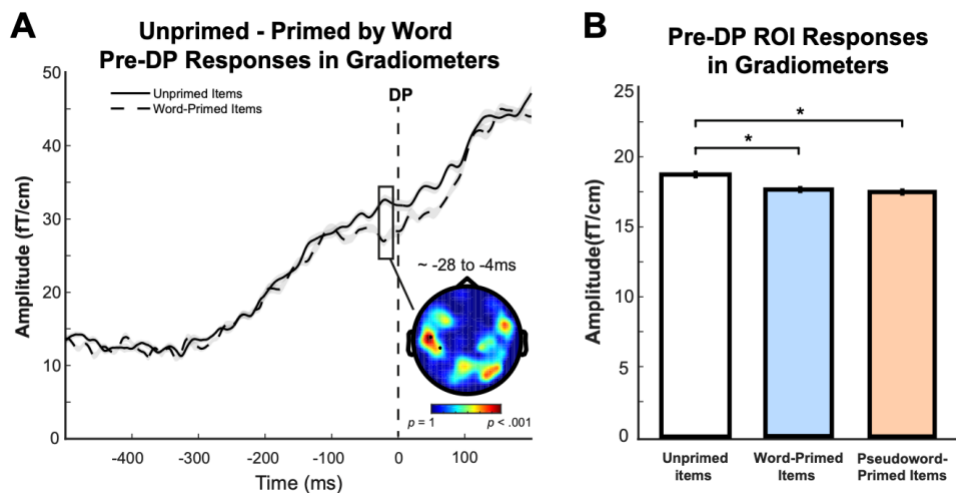


Figure 2.5 Pre-DP results. **A.** Pre-DP response difference between items that are unprimed and primed by a word in MEG gradiometer sensors within -150 to 0ms (a time window at which words and pseudowords are indistinguishable). The topographic plot shows F-statistics for the entire sensor array with the scalp locations that form a statistically significant cluster

highlighted and marked with black dots. Waveforms represent MEG response averaged over the spatial extent of the significant cluster shown in the topography. The grey shade of waveforms represents \pm within-participant SE, adjusted to remove between-participants variance (Cousineau, 2005). **B.** ROI analysis of neural responses evoked by unprimed and primed items averaged over the same pre-DP time period of -150-0ms but across gradiometer sensor locations which showed the post-DP pseudoword>word lexicality effect (see Figure 2.6A). Bars are color-coded by prime type on the x axis (unprimed items, no fill; word-primed items, blue; pseudoword-primed items, orange; black frame indicates that words and pseudowords are indistinguishable). All error bars represent \pm within-participant SE, adjusted to remove between-participant variance. Statistical significance: * $p < 0.05$.

Post-DP analyses. We then examined the post-DP response differences between words and pseudowords (lexicality effect). The gradiometer sensors showed a significant cluster of 39335 sensor x time points ($p < .001$) over the left side of the scalp at 313-956ms post-DP (Figure 2.6A). In this cluster, pseudowords evoked a significantly stronger neural response than words. Similarly, magnetometer sensors also detected a significant left-hemisphere cluster of 68517 sensor x time points ($p < .001$) at 359-990ms post-DP (Figure 2.6B) showing the same lexicality effect. We did not find any significant cluster in which words evoked greater neural responses than pseudowords. These results are consistent with findings from Gagnepain and colleagues (2012). To locate the likely neural source of the effects found in sensor space, we conducted source reconstruction by integrating gradiometers and magnetometers. As shown in Figure 2.6C, results from source space showed that neural generators of the lexicality effect were estimated to lie within the left superior temporal gyrus (STG, volume of 2315 voxels, $p < .001$, peak at $x = -46$, $y = -36$, $z = 0$; $x = -52$, $y = -34$, $z = -6$; $x = -56$, $y = -28$, $z = -10$). This

location, and direction of response, is consistent with a sub-lexical (e.g. phonemic) process being modulated by lexicality; in line with the predictive-selection account.

Next, we investigated whether the neural responses that were modulated by lexicality were also influenced by prime type by conducting an ROI analysis which tested the interaction between prime type and lexicality, as well as planned pairwise comparisons of priming effects on words alone, using data averaged over the time window and the sensor locations of the significant cluster shown in panel A and B (Figure 2.6D & E). Since these planned pairwise comparisons involve responses to familiar words only (i.e. words that are word-primed vs unprimed, words that are word-primed vs pseudoword-primed), they are orthogonal to the lexicality effect that defined the pseudoword>word cluster and hence are not confounded by task. The interaction was significant in both gradiometers, $F(1.96, 41.11) = 7.30, p = .002$, and magnetometers, $F(1.90, 39.99) = 5.80, p = .007$. Specifically, there was a significant effect of prime type for words, $F(1.93, 40.55) = 8.01, p = .001$ (gradiometers), $F(1.81, 37.96) = 5.61, p = .009$ (magnetometers), such that neural signals evoked by word-primed words were significantly stronger than those evoked by unprimed words, $t(21) = 2.22, p = .019$ (gradiometers), $t(21) = 3.33, p = .002$ (magnetometers), and pseudoword-primed words, $t(21) = 3.70, p < .001$ (gradiometers), $t(21) = 2.64, p = .008$ (magnetometers). In contrast, there was no reliable main effect of prime type for pseudowords, $F(1.94, 40.80) = 0.67, p = .514$ (gradiometers), $F(1.79, 37.61) = 0.80, p = .446$ (magnetometers).

The corresponding tests performed on the source-reconstructed power within the lexicality ROI of suprathreshold voxels (Figure 2.6F) did not show a reliable interaction effect between lexicality and competitor priming, $F(1.56, 32.85) = 0.99, p = .360$. Nevertheless, consistent with sensor space results, source power indicated a significant effect of prime type for words, $F(1.73, 36.42) = 3.77, p = .038$, but not pseudowords, $F(1.62, 33.94) = 1.12, p = .326$. Pairwise comparisons also indicated that word-primed words evoked significantly greater

source strength than unprimed words, $t(21) = 2.66$, $p = .007$, though the effect between word-primed and pseudoword-primed words was not significant, $t(21) = 1.26$, $p = .110$. Overall, in line with behavioural results, neural responses evoked by words and pseudowords were also influenced differently by prime type. Critically, competitor priming modulated the post-DP neural responses evoked by words, but not those evoked by pseudowords, and these effects were localised to the left STG regions that plausibly contribute to sub-lexical processing of speech. This matches the pattern of responses proposed in the predictive-selection model (see Figure 2.1F).

As encouraged by a reviewer, we also conducted whole brain analyses for the competitor priming effects. We found a significant word-primed word > unprimed word cluster of 1197 sensor x time points ($p = .034$) in magnetometers in the left hemisphere within a time window of 426 - 466ms post-DP. We also found a significant and a marginal word-primed word > pseudoword-primed word cluster in gradiometers in the left hemisphere respectively of 527 sensor x time points ($p = .011$) at 719-749ms and 471 sensor x time points ($p = .053$) at 315-336ms. These topographies and time courses overlap with the pseudoword > word clusters and are consistent with our ROI results. Hence, the ROI analyses have picked up the most important findings from these whole-brain analyses.

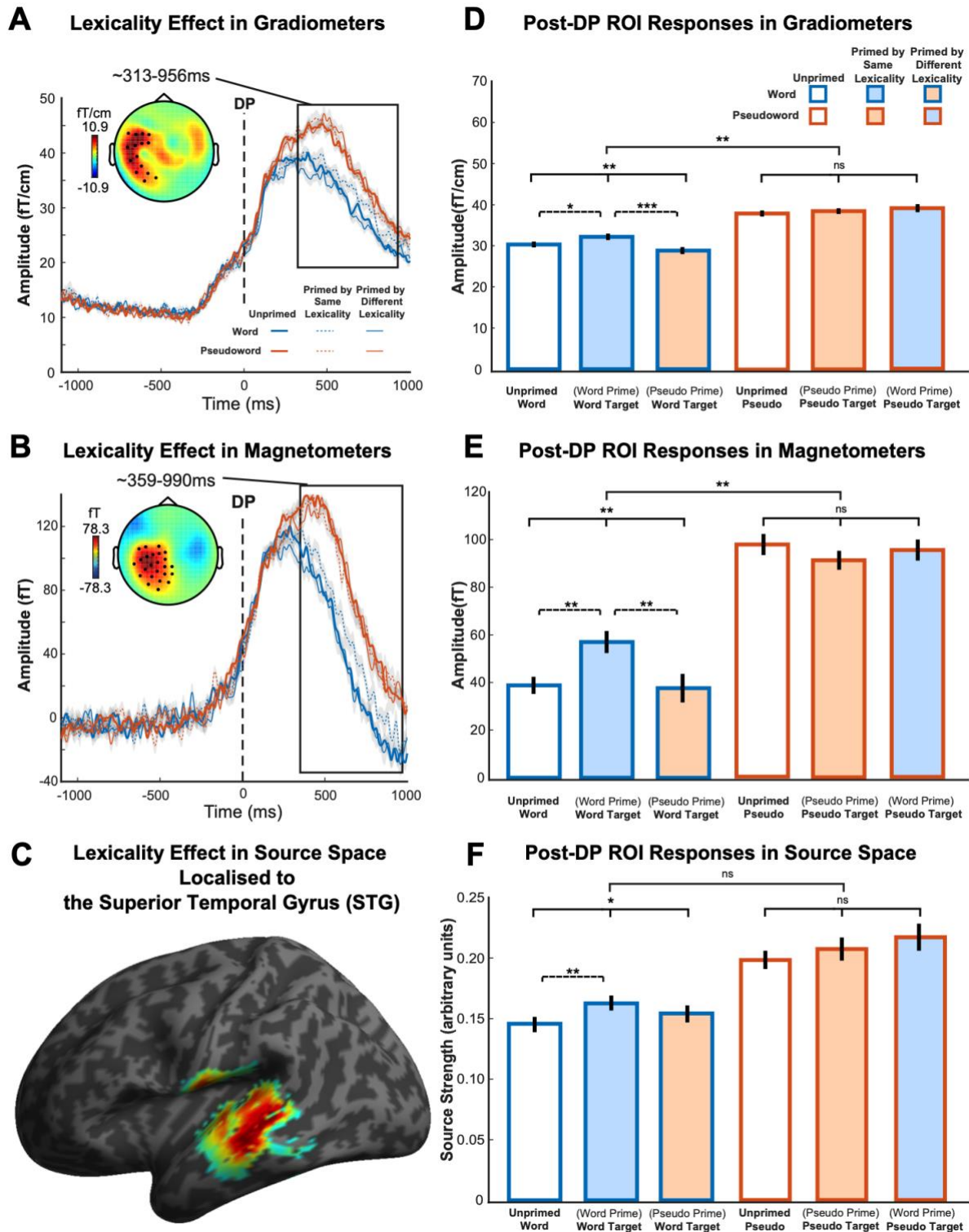


Figure 2.6 Post-DP results showing lexicality effects and corresponding ROI responses evoked by conditions of interest. **A & B.** Post-DP lexicality effects in MEG gradiometer and magnetometer sensors. The topographic plots show the statistically significant cluster with a main effect of lexicality (pseudoword > word). Waveforms represent MEG response averaged

over the spatial extent of the significant cluster shown in the topography. The grey shade of waveforms represents \pm within-participant SE, adjusted to remove between-participants variance. **C.** Statistical parametric map showing the cluster (pseudoword > word) rendered onto an inflated cortical surface of the Montreal Neurological Institute (MNI) standard brain thresholded at FWE-corrected cluster-level $p < 0.05$, localised to the left STG. **D, E & F.** Post-DP ROI ANOVA on neural signals and source strength evoked by conditions of interest averaged over the time window and scalp locations of the significant cluster shown in panel A, B & C. Bars are color-coded by lexicality and prime type on the x axis (words, blue frame; pseudowords, orange frame; unprimed, no fill; primed by same lexicality, consistent fill and frame colors; primed by different lexicality, inconsistent fill and frame colors). All error bars represent \pm within-participant SE, adjusted to remove between-participants variance. Statistical significance from ANOVAs: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Statistical comparisons shown with solid lines indicate the lexicality by prime-type interaction and main effects of prime-type for each lexicality, whereas comparisons with broken lines indicate the significance of planned pairwise comparisons.

To ensure that other response patterns were not overlooked, we also investigated whether there was any lexicality by prime-type interaction at other locations across the scalp and source spaces, and during other time periods. As shown in Figure 2.7A, a significant cluster of Gradiometers at midline posterior scalp locations were found at 397-437ms post-DP, in which the effect of priming was significantly different for words and pseudowords. Figure 2.7B shows gradiometer signals evoked by conditions of interest averaged over the spatial and temporal extent of the significant cluster in panel A. To explore this profile, we computed an orthogonal contrast to assess the overall lexicality effect (the difference between words and pseudowords),

and the result was marginal, $F(1.00, 21.00) = 3.50$, $p = .075$. The effect of prime type was marginally significant for words, $F(1.89, 39.78) = 3.08$, $p = .060$, but significant for pseudowords, $F(1.80, 37.85) = 7.14$, $p = .003$. The location and pattern of this interaction cluster were dissimilar to those predicted by either competitive- or predictive-selection theories and no cluster survived correction in magnetometer sensors or source space hence we did not consider this effect to be as relevant or interpretable as our other findings. I report it here in the interest of completeness and transparency.

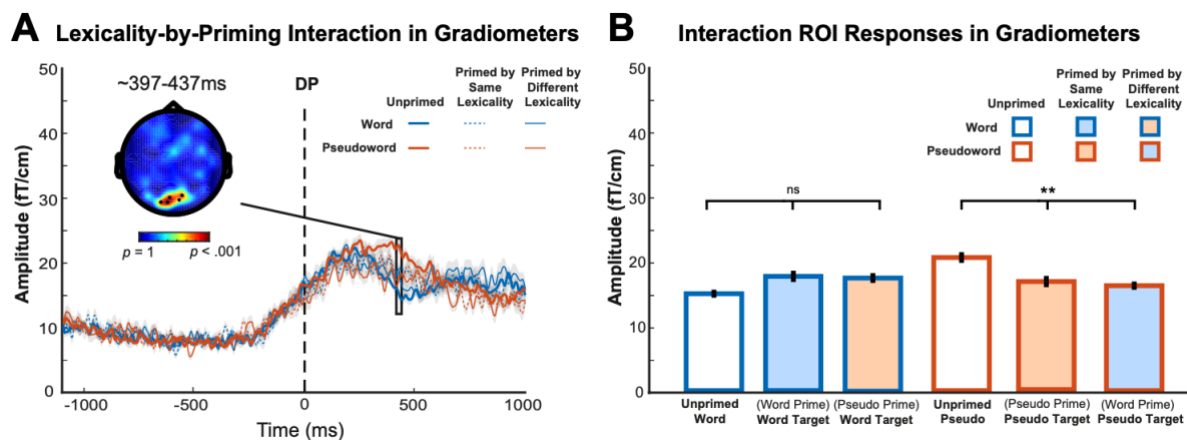


Figure 2.7 Post-DP results showing lexicity-by-priming interaction effects in MEG gradiometers. **A.** The topographic plot shows F -statistics for the statistically significant cluster that showed an interaction between lexicity and prime type. Waveforms represent gradiometer responses averaged over the spatial extent of the significant cluster shown in the topography. The grey shade of waveforms represents \pm within-participant SE, adjusted to remove between-participants variance. **B.** Gradiometer signals evoked by conditions of interest averaged over temporal and spatial extent of the significant cluster in panel A. All error bars represent \pm within-participant SE, adjusted to remove between-participants variance. Statistical significance: ** $p < 0.01$. The statistical comparison lines indicate main effects of prime type for each lexicity. The lexicity by prime-type interaction is statistically reliable as expected based on the defined cluster.

Linking neural and behavioural effects. To further examine the relationship between neural and behavioural response differences attributable to competitor priming or lexicality, we conducted a single-trial regression analyses using linear mixed-effect models that account for random intercepts and slopes for participants and stimuli sets (grouped by their initial segments). We calculated behavioural RT differences and neural MEG differences caused by: (1) lexicality. i.e. the difference between pseudoword and word trials (collapsed over primed and unprimed conditions) and (2) competitor priming, i.e. the difference between unprimed and word-primed word trials, with MEG signals averaged over the spatial and temporal extent of the post-DP pseudoword>word cluster seen in sensor space and the STG peak voxel in source space (see Figure 2.6). We then assessed the relationship between these behavioural and neural difference effects in linear mixed-effect regression of single trials, with differences in RTs as the independent variable and differences in MEG responses as the dependent variable. The analyses were conducted using the lme4 package in R (Bates et al. 2014).

As shown in Figure 2.8A, we observed a significant positive relationship between RTs and magnetometers on lexicality difference ($\beta = 0.11$, $SE = 0.01$, $t(23.31) = 7.77$, $p < .001$), although associations between RTs and gradiometers or source response were not significant. These observations from magnetometers indicated that slower lexical decision times evoked by pseudowords were associated with greater neural responses. Furthermore, the intercept parameter for the magnetometers model was significantly larger than zero, $\beta = 37.58$, $SE = 5.72$, $t(23.09) = 6.57$, $p < .001$. We can interpret this intercept as the neural difference that would be predicted for trials in which there was no delayed response to pseudowords compared to words. The significant intercept indicated a baseline difference in neural responses to words and pseudowords, even in the absence of any difference in processing effort (as indexed by

lexical decision RTs). This suggested the engagement of additional neural processes specific to pseudowords regardless of the behavioural effect (cf. Taylor et al., 2014).

Figure 2.8B showed another significant positive relationship between RTs and magnetometers on competitor priming difference ($\beta = 0.15$, $SE = 0.02$, $t(38.85) = 7.89$, $p < .001$), while relationships between RTs and gradiometers or source response were again not significant. Interestingly, unlike for the lexicality effect, the intercept in this competitor priming magnetometers model did not reach significance ($\beta = 12.88$, $SE = 7.27$, $t(21.33) = 1.77$, $p = .091$). This non-significant intercept might suggest that if word-primed words did not evoke longer RTs than unprimed words, magnetometer signals would not be reliably different between the two conditions either. Hence, consistent with predictive-selection accounts, the increased post-DP neural responses in the STG caused by competitor priming was both positively linked to and mediated by longer response times.

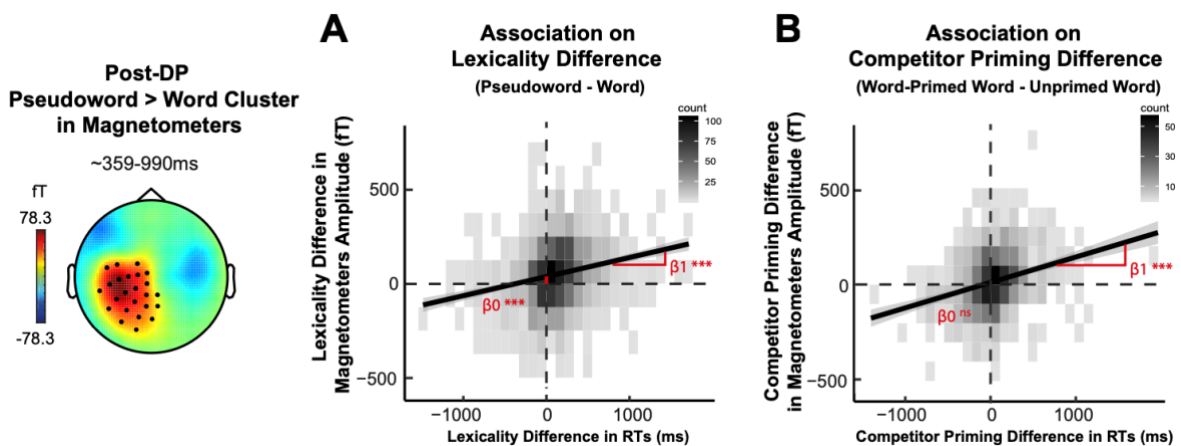


Figure 2.8 Single-trial linear mixed-effect models which accounted for random intercepts and slopes for participants and stimuli sets (grouped by initial segments) were constructed to compute the relationship between RTs and magnetometers on (A) lexicality difference (i.e. between pseudowords and words, collapsed over unprimed and primed conditions) and (B) competitor priming difference (i.e. between word-primed word and unprimed word conditions). Magnetometer responses were averaged over the time window and scalp locations of the

significant post-DP pseudoword>word cluster (see Figure 2.6). β_1 refers to the model slope, β_0 refers to the model intercept. Statistical significance: *** $p < 0.001$.

2.4 Discussion

In this Chapter, I reported an MEG study which distinguished different implementations of Bayesian perceptual inference by manipulating the prior probability of spoken words and examining the pattern of neural responses. We replicated the competitor priming effect such that a single prior presentation of a competitor word (e.g. *hijack*) delayed the recognition of a similar-sounding word (e.g. *hygiene*), whereas this effect was not observed when the prime or target was a pseudoword (e.g. *hijure*). Armed with this behavioural evidence, we used MEG data to test the neural bases of two Bayesian theories of spoken word recognition.

2.4.1 Competitive- vs predictive-selection

Competitive-selection accounts propose that word recognition is achieved through direct inhibitory connections between representations of similar candidates (e.g. McClelland & Elman, 1986). Priming boosts the activation of heard words and increases lateral inhibition applied to neighbouring words, which delays their subsequent identification. The effect of competitor priming is to increase lexical uncertainty, and hence lexical-level neural responses, until later time points when target words can be distinguished from the competitor prime (Figure 2.1C). In contrast, predictive-selection accounts propose that word recognition is achieved by subtracting predicted speech from heard speech and using computations of prediction error to update lexical probabilities (Davis & Sohoglu, 2020). By this view,

predictions for segments that are shared between competitor primes and targets (pre-DP segments) will be enhanced after presentation of prime words. Thus, competitor priming will reduce the magnitude of prediction error, and hence neural responses pre-DP (Figure 2.1F). Only when speech diverges from predictions (post-DP segments) will competitor-primed words evoke greater prediction error, leading to increased neural response in brain areas involved in pre-lexical (e.g. phonemic) processing of speech (Blank et al., 2018; Blank & Davis, 2016).

It should be acknowledged that both models involve multiple levels of representation and hence both sub-lexical and lexical processes. However, our focus is on lexical processing within the competitive-selection framework and sub-lexical processing within the predictive-selection framework. These are the critical levels that 1) support word recognition according to each theory, 2) are modulated by the competitor priming effect that our study manipulates and 3) are invoked to explain the slower behavioural responses and associated changes in MEG responses that we observed.

We tested the predictions for the direction and timing of neural responses associated with competitor priming using MEG data which showed opposite neural effects pre- and post-DP. In the pre-DP period, consistent with predictive-selection but contrary to competitive-selection mechanism, we saw decreased neural responses for word-primed items compared to unprimed items. The initial, shared segments between prime (*hijack*) and target (*hygiene*) words evoked a reduced response during early time periods in line with a reduction in prediction error. However, during the post-DP period, we found competitor-primed words evoked stronger neural responses than unprimed words in exactly the same locations and time periods that showed increased responses to pseudowords (*hijure*) compared to words. These post-DP response increases are in line with enhanced processing difficulty for competitor-primed words and pseudowords due to greater prediction error. Thus, the time course of the competitor

priming neural effects – showing reduced neural responses pre-DP and increased neural responses post-DP – closely resembles the expected changes in prediction error (Figure 2.1F) based on predictive-selection mechanisms.

On top of the direction and timing of neural responses, effects of lexicality and competitor priming localised to the left STG. This is a brain region that has long been associated with lower-level sensory processing of speech (Yi et al., 2019). Our observation of increased responses to pseudowords in this region is in accordance with source-localised MEG findings (Gagnepain et al., 2012; Shtyrov et al., 2012) and evidence from a meta-analysis of PET and fMRI studies (Davis & Gaskell, 2009). This location is also consistent with the proposal that lexical influences on segment-level computations (rather than lexical-level computations themselves) produce reliable neural differences between words and pseudowords (Davis & Sohoglu, 2020). We take this finding as further evidence in favour of computations of segment prediction error as a critical mechanism underlying word identification. Increased prediction error for pseudowords has also been linked to encoding of novel lexical items in theoretical work (Davis & Sohoglu, 2020) and in studies of word learning in young children (Ylinen, et al 2017).

We further show using regression analyses that neural (MEG) and behavioural (RT) effects of lexicality and competitor priming are linked on a trial-by-trial basis. Trials in which pseudoword processing or competitor priming leads to larger increases in RT also have greater post-DP neural responses. These links between behavioural and neural effects of lexicality and competitor priming are once more in-line with the proposal that post-DP increases in prediction error are a key neural mechanism for word and pseudoword processing and can explain the delayed behavioural responses seen in competitor priming. Interestingly, lexicality and competitor priming effects differ in terms of whether a reliable neural response difference would be seen for trials with no baseline RT difference. While neural lexicality effects were

significant even for trials that did not show behavioural effects, the same was not true for the competitor priming effect. These results indicate that, consistent with predictive-selection accounts, the post-DP neural competitor priming effect was mediated by changes in behavioural RTs. Only those trials in which competitor priming slowed behavioural responses led to larger neural responses. In contrast, an increased neural response to pseudowords was expected even in trials for which RTs did not differ between pseudowords and words. We will consider the implications of these and other findings for pseudoword processing in the next section.

2.4.2 How do listeners process pseudowords?

Participants identified pseudowords with a speed and accuracy similar to that seen during recognition of familiar words. This is consistent with an optimally-efficient language processing system (Marslen-Wilson, 1984; Zhuang et al, 2014), in which pseudowords can be distinguished from real words as soon as deviating speech segments are heard. Beyond this well-established behavioural finding, however, we reported two seemingly contradictory observations concerning pseudoword processing.

The first is that, while post-DP neural activity and response times for words were modulated by competitor priming, processing of pseudowords was not similarly affected. This might suggest that the prior probability of hearing a pseudoword and the prediction error elicited by mismatching segments are not changed by our experimental manipulations. This may be because pseudowords have a low or zero prior probability and elicit maximal prediction errors that cannot be modified by a single prime. Yet, memory studies suggest that even a single presentation of a pseudoword can be sufficient for listeners to establish a lasting memory trace (Mckone & Trynes, 1999; Arndt et al., 2008). However, it is possible that this memory for

pseudowords reflects a different type of memory (e.g. episodic memory) from that produced by a word, with only the latter able to temporarily modify long-term, lexical-level representations and predictions for word speech segments (cf. Complementary Learning Systems theories, McClelland et al., 1995; Davis & Gaskell, 2009). Additionally, these differences between words and pseudowords may be influenced by the lexical decision task, which may have implicitly cued participants to treat words and pseudowords differently. Participants need to identify the exact form of a single word in order to confirm its lexical status, but a deviation from all known words needs to be established to indicate a pseudoword (Norris & Kinoshita, 2008).

A second observation is that, contrary to the null result for post-DP processing, pseudoword priming reduced subsequent pre-DP neural responses evoked by target items to a similar degree as word priming (Figure 2.5B). This pre-DP effect is surprising given previous evidence suggesting that pseudowords must be encoded into memory and subject to overnight, sleep-associated consolidation in order to modulate the speed of lexical processing (Tamminen et al., 2010; James et al., 2017) or neural responses (Davis & Gaskell, 2009; Landi et al. 2018). It might be that neural effects seen for these pre-DP segments were due to changes to the representation of familiar words that our pseudowords resembled, though these were insufficient to modulate processing of post-DP segments.

2.4.3 Conclusion

This chapter described an MEG study that provided compelling evidence in favour of neural computations of prediction error during spoken word recognition. Although the previous work by Gagnepain et al. (2012) also provided evidence for the predictive-selection account, their behavioural effects of consolidation on word recognition were obtained during different tasks and different sessions from their neural responses. This study goes beyond the previous work

by adopting a single task (lexical decision) and using a competitor priming paradigm that permits concurrent measurement of perceptual outcomes and neural responses in a single session. This enables us to directly link trials that evoked stronger neural signals in the STG to delayed RTs and hence provide stronger evidence that both of these effects are caused by competitor priming.

In addition, unlike previous work (Brodbeck et al. 2018; Donhauser & Baillet, 2020) which reported neural responses correlated with lexical entropy as well as prediction error (surprisal), we did not find any similarly equivocal evidence. These earlier studies measured neural responses to familiar words in continuous speech sequences such as stories or talks. It might be that effects of lexical entropy are more apparent for connected speech than isolated words. However, since lexical uncertainty (entropy) and segment-level predictability (segment prediction error or surprisal) are highly correlated in natural continuous speech, these studies may be less able to distinguish between the lexical and segmental mechanisms that we assessed here. In contrast, our speech materials were carefully selected to change lexical probability (through priming) and for priming to have opposite effects on segment prediction error before and after DP. This manipulation provides evidence in favour of predictive-selection mechanisms that operate using computations of prediction error during spoken word recognition.

3 LEXICAL AND SUB-LEXICAL PROCESSING IN SPOKEN WORD RECOGNITION

3.1 Introduction

The MEG study reported in Chapter 2 confirmed that the competitor priming effect involves lexical processing and is not affected by non-lexical knowledge (e.g. episodic or other source of memory for recently presented pseudowords), since it was seen for word-primed words only but not for when pseudowords served as the prime or the target. Indeed, Dufour and Nguyen (2017) also showed that the competitor priming effect was not modulated by talker differences between prime and target items when the prime was presented once only. In addition, the robust effect lasting across separate blocks of stimuli was evoked by just a single prior presentation of a competitor word. This is in line with Bayesian mechanisms which propose probabilistic processing of words while also indicating strong cognitive flexibility of the human brain. The MEG results from the previous chapter also showed that the differential neural responses to words and pseudowords during competitor priming were manifested at the same sub-lexical level, which is consistent with top-down lexical predictions. In this chapter, I will further investigate whether the changes in lexical prediction that gives rise to competitor priming continue to be observed even when lexical processing is not required for task performance.

Competitor priming as supported by models of spoken word recognition

The behavioural effect of competitor priming - delayed word recognition caused by prior presentation of a competitor word beginning with the same initial segments - is supported by major spoken word recognition models although through different mechanisms.

In localist models such as TRACE (McClelland & Elman, 1986; Figure 3.1A), lexicons are represented in separate nodes and selection of word candidates is achieved through parallel process of competition. Its hierarchical structure allows facilitative interactions between nodes in different levels (i.e. features, phonemes and words levels) and inhibitory competitions within the same representational level. Input from the phonological and feature levels are fed onto the lexical level, which drives word recognition. In TRACE, word frequency (i.e. prior probability of words in Bayesian term) not only determines the baseline activation level of each lexical unit, but also changes the weights between phoneme and word nodes. Hence, lexical frequency plays a role in recognition as soon as word candidates become active (Dahan et al., 2001). During competitor priming, prior presentation of the competitor word enhances its frequency, hence it will be more activated when the initial segments of the target word is heard and consequently delays the recognition of the target word due to lateral inhibition.

While TRACE is a hierarchical model that include lexical, phoneme and feature levels, the effect of phonotactic probability (i.e. the likelihood of phoneme sequence) is only reflected through interactions with the lexical level and does not affect recognition of words directly (Vitevitch & Luce, 1998). When a lexical competitor is presented and activated, the phonotactic probability of the pre-DP segments shared with the target word should be enhanced. Therefore, contrary to the inhibitory lexical competitor priming effect, the processing of the coherent pre-DP phonological segment should lead to facilitatory recognition. However, whether these different effects of lexical and sub-lexical processing can be simulated by TRACE is unclear.

In the Predictive Coding account (PC; Davis & Sohoglu, 2020; Figure 3.1B), calculations of prediction error – by subtracting top-down predictions from bottom-up speech input – drives word recognition. This account proposes that word frequency modulates connection weights at the phoneme level. Predictions of the next segments are strengthened by

higher connection weights, while prediction error reflects the disruption of such connection. Hence, during competitor priming, prior presentation of the competitor word strengthens the connection weights between pre-DP and post-DP segments of the prime item, but inhibits the segment connection for the target word, which delays its recognition.

At the same time, facilitatory phonological priming effect is also incorporated by the PC. This is because the prediction for the phoneme sequence before the DP is strengthened during the presentation of the competitor prime, and this prediction matched with the input of the target pre-DP segment, hence prediction error for this sound sequence should be reduced. These hypotheses were also confirmed by the reduced pre-DP neural results of competitor primed items as described in Chapter 2.

In addition, the Distributed Cohort Model (DCM; Gaskell & Marslen-Wilson, 1997; Figure 3.1C) proposes an architecture in which phonological representations and lexical representations of all words are blended on the same nodes that map directly from the low-level feature input. Hence, competition is an intrinsic part of the model due to the interference of multiple distributed representations. The size of candidates and their relative frequency strongly affect semantic activation. However, unlike the relatively fixed property of lateral inhibition in TRACE, the DCM proposes that the effects of competition vary across lexical and phonological dimensions of the distributed representational space. As shown in cross-model studies by Gaskell and Marslen-Wilson (1999, 2002), interference was strong at the semantic level between competitive words, but phonological coherence between cohort members reduced interference and allowed for facilitatory activation when the priming segment was incomplete and ambiguous (i.e. matching multiple activated lexical items).

In the competitor priming paradigm, words from the same cohort (e.g. *hygiene* and *hijack*) are coherent in terms of the phonological features before the deviation point but not after, and compete with each other constantly on semantic nodes. Higher frequency of the

prime word, due to its prior presentation, would indicate a larger distance between the output of the network (target word) and the expected lexical representations (closer to prime word), hence leading to lower activation of the target word, which delays its recognition. However, due to the blended representations of phonology and meaning, the effects of both lexical and sub-lexical processing may be reflected through suitable tasks and stimuli manipulations.

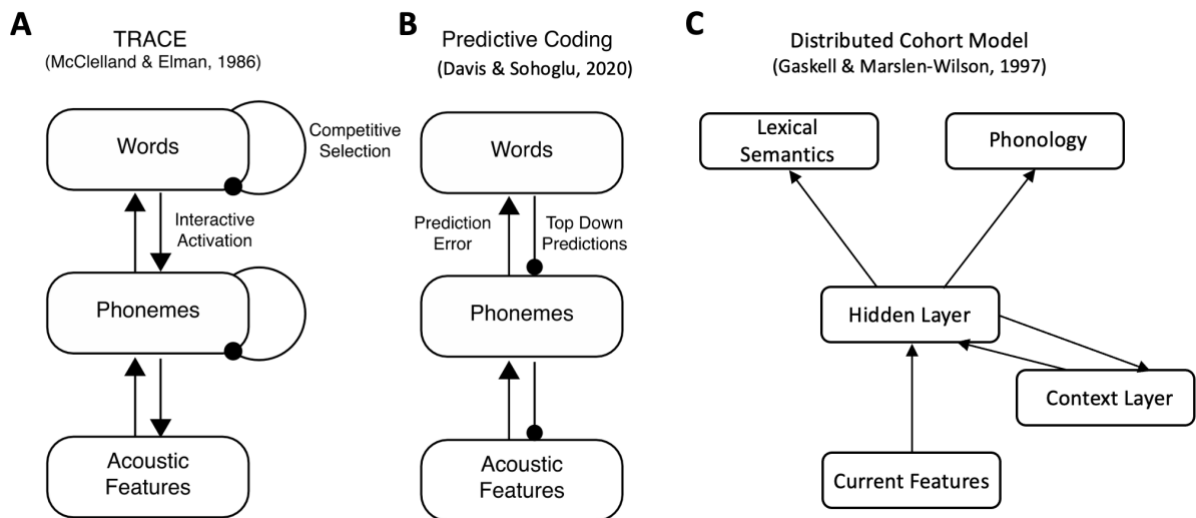


Figure 3.1. Hierarchical structure of spoken word recognition models: TRACE (McClelland & Elman, 1986), Predictive Coding (Davis & Sohoglu, 2020) and Distributed Cohort Model (Gaskell & Marslen-Wilson, 1997).

Motivations and predictions for studies reported in this chapter

In order to probe both lexical and phonological processing during spoken word recognition, we conducted a series of experiments that adopted the lexical decision task, which involves explicit lexical judgment, as well as the pause detection task (Davis et al., 2009; Dumay & Gaskell 2007; Gagnepain et al. 2012; Gaskell & Dumay, 2003; Mattys & Clark, 2002), which leads participants to focus on the phonetic features of the stimuli while measures the overall lexical activity implicitly at the location of the pause at the same time. The use of the pause detection

task also serves the purpose of evoking the same kind of responses in words and pseudowords. This contrasts with the lexical decision task, which always triggers a ‘yes’ response on words and a ‘no’ response on pseudowords, hence potentially evokes different cognitive processes on these two types of stimuli (Norris & Kinoshita, 2008). Additionally, inserting a pause in the stimulus will also temporarily separate the stimulus into two independent phonological segments, hence disrupts the completeness of the word. The manipulation of word completeness (Gaskell & Marslen-Wilson, 2002) together with the pause detection task can help dissociate effects of phonological processing from lexical processing.

To test whether the competitor priming effect observed during lexical processing can also be observed when the task focuses on phonological processing, we used the lexical decision task in Experiment 2 and the pause detection task in Experiment 3 respectively on a similar competitor priming paradigm as the one adopted by the MEG study (see Chapter 2). In order to simplify the results interpretation, we kept all prime and target words pause absent in Experiment 3 (while using pause present filler words to balance the responses). All three accounts discussed above predict that the competitor priming effect should be observed in both experiments. In particular, although these models account for lexical and sub-lexical processing during competitor priming, the completeness of the word representations also matters. Since the competitor word primes were presented in full and without any disruptions from pauses in both experiments, any initial phonological coherence evoked between the cohort prime and target words should not last after the DP while the semantic competition effect persists, hence facilitatory effect from phonological processing are not likely to be observed in behavioural RT results.

To further dissociate phonological and lexical processing, we adopted a study design with separate prime and target phases in Experiment 4a, 4b, 5a and 5b. Lexical decision task and pause detection task were mixed in the two phases (for specific tasks used in each

experiment, see Table 3.1). This way, the effect of task switching in a design with interleaved prime and target items can be avoided. Moreover, in order to increase the power, all prime items were presented four times, in particular, twice with pause and twice without pause during the pause detections task. Through this design, lexical decision and pause detection tasks used during the prime phase could help dissociate the effect of inhibitory lexical processing caused by presentations of competitor primes and the effect of facilitatory sub-lexical processing caused by presentations of ambiguous phonological segments (temporarily isolated by pause) shared by both prime and target words. In addition, the different tasks used during the target phase aimed to test whether these effects (modulated by frequency changes at either lexical or sub-lexical level) is also affected by task-evoked decisions as suggested by Balota and Chumbley (1984). While the opposite effects of lexical and sub-lexical processing are embodied in TRACE, facilitatory phonological-level effects are masked by the inhibitory lexical-level effects in this model. In contrast, these opposite effects could be accounted for by computations of prediction error at the phoneme level in the PC account and the distributed structure of the DCM.

Table 3.1. Experiments and corresponding tasks

Experiments	Tasks
Experiment 2	Lexical Decision (LD)
Experiment 3	Pause Detection (PD)
Experiment 4a	Prime phase: PD – Target phase: LD
Experiment 4b	Prime phase: LD - Target phase: LD
Experiment 5a	Prime phase: PD - Target phase: PD
Experiment 5b	Prime phase: LD - Target phase: PD

3.2 Experiment 2

Experiment 2 used the same lexical decision task as Experiment 1 (the MEG study) for replication purposes. In order to further examine whether the competitor priming effect is unique to word-primed word, we added another two conditions that involved words primed by two similar pseudowords and pseudowords primed by two similar words to test if double priming by items of different lexicality would evoke any stronger effect than a single prime.

3.2.1 Method

Participants

Thirty-three native British English speakers aged between 18 to 35 were tested on this experiment. Among which, sixteen participants were recruited from the volunteer panel of the MRC Cognition and Brain Sciences Unit and were tested in the lab, while the rest were recruited through a website platform, Prolific Academic (www.prolific.ac), and were tested online. Online testing largely accelerated the speed of data collection and became especially helpful due to the pandemic (although the data for this study was collected pre-pandemic), while lab testing ensured the quality of the research procedure and provided a reference for the data collected through online testing. All recruited participants provided informed consent under a process approved by the Cambridge Psychology Research Ethics Committee and received monetary compensation. None of them had any history of neurological disorder or hearing impairment based on self-report. One participant recruited online was removed from subsequent analyses due to high error rate (two standard deviations above the mean), resulting in 32 participants in total.

Experimental Design and Materials

The experiment was implemented in JavaScript using JSpsych (de Leeuw, 2015) so that it can be presented in web browsers for both lab and online testing. The data were collected through the web server tool JATOS (Lange et al., 2015). This experiment was also built on the same lexicality (word, pseudoword) and prime type factors (unprimed, primed by same lexical status, prime by different lexical status) from Experiment 1. However, here we extended the paradigm of Experiment 1 by adding an additional prime type level: double-primed by different lexical status, resulting in two additional conditions: two pseudowords-primed word and two words-primed pseudoword, on top of the original single word-primed word, single word-primed pseudoword, single pseudoword-primed word and single pseudoword-primed pseudoword conditions (Figure 3.2). The same lexical decision task was used as in experiment 1. Participants were asked to press a key on the keyboard with their right index finger if they heard a word and another key with their left index finger if they heard a pseudoword. 480 unique spoken items were presented in 4 blocks of 120 trials, with each block lasting approximately 5 minutes. Like before, each prime-target pair was separated by ~20-80 trials of items and stimuli were Latin-square counterbalanced across participants and pseudo-randomised within each block.

We used the same set of spoken stimuli as those for the MEG study described in Chapter 2, with stimuli in each of the 160 quadruplet sets containing two words and two pseudowords sharing the same initial segments and diverging only after the deviation point (DP). However, we used the original stimulus recordings here rather than morphing the pre-DP segments of the 4 stimuli within each quadruplet set.

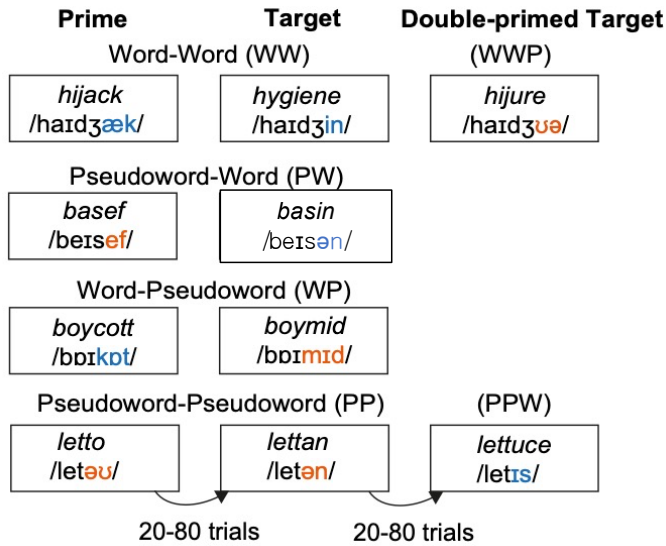


Figure 3.2. The design of Experiment 2 showing all conditions. On top of the original conditions used in the MEG study – word-primed word (WW), pseudoword-primed word (PW), word-primed pseudoword (WP) and pseudoword-primed pseudoword (PP), this experiment added two more conditions – double word-primed pseudoword (WWP) and double pseudoword-primed word (PPW). All prime and target pairs share the same initial segments, which are coloured in black in the example phonetic transcriptions, words are shown with blue ending and pseudowords are shown with orange ending in the transcriptions.

Data analysis

The response times (RTs) were measured from the DP of the stimuli, since it is impossible to make accurate lexical judgements before the deviation point. The RTs were then transformed based on the normality of Q-Q plots and histograms (i.e. which transformation made the data and residuals more normally distributed) and we chose the transformation that maximises the normality of the data and residuals. In accordance with this principle, we log transformed the RTs, although all Figures report untransformed response times for clarity. See Appendix C for descriptive statistics for the data presented in all Figures in this Chapter.

All data were analysed using mixed-effect models, with lexicality and prime type as fixed factors and participants and items as random factors. Transformed RTs and error rates were respectively analysed using linear and logistic mixed-effect models using the *lme4* package implemented in R (Bates et al. 2014). Like in Experiment 1, maximal models accounting for all random effects were attempted wherever possible, but reduced random effects structures were applied when the full model did not converge (Barr et al., 2013). Likelihood-ratio tests comparing the full model to a nested reduced model using the Chi-Square distribution were conducted to evaluate the main effect of lexicality and the interaction between lexicality and prime type. Significance of individual model coefficients were obtained using t (reported by linear mixed-effect models) or z (reported by logistic mixed-effect models) statistics from the model summary. One-tailed t statistics for RTs are reported for the following planned pairwise contrasts: (1) word-primed words versus unprimed words, (2) word-primed words versus pseudoword-primed words.

When assessing priming effects, we excluded data from target trials in which the participant made an error in the corresponding prime trial, because it is unclear whether such target items will be affected by priming given that the prime word was not correctly identified (Monsell & Hirsh, 1998). In addition, trials with RTs shorter than 200ms were removed from further analysis, since responses shorter than 200ms are quicker than the minimal time required for stimulus processing (Woods et al., 2015).

3.2.2 Results

Response Times. Figure 3.3A shows results of factorial analysis of lexicality (word, pseudoword) and prime type (unprimed, primed by same lexical status, primed by different lexical status and double primed by different lexical status). The plot reports untransformed

results for clarity (see Appendix C for descriptive statistics for the data presented in the figure). Consistent with the behavioural results from the MEG Study, there was a significant main effect of lexicality, in which RTs for pseudowords were significantly longer than for words, $X^2(3) = 35.72, p < .001$, and a significant interaction between lexicality and prime type, $X^2(2) = 16.05, p < .001$. We then conducted two separate one-way models for words and pseudowords respectively, with prime type as the only factor. The word model showed a significant effect of prime type, $X^2(3) = 12.42, p = .006$, while the pseudoword model showed a marginal effect, $X^2(3) = 7.55, p = .056$. Our planned pairwise contrasts again showed the competitor priming effect, i.e. word-primed words were recognised significantly more slowly than unprimed words (results reported here are in log scale, for mean raw RTs see Fig 3.2A), $\beta = 0.05, SE = 0.01, t(31.73) = 3.72, p < .001$, and single pseudoword-primed words, $\beta = 0.03, SE = 0.01, t(82.68) = 2.86, p = .003$. After Tukey correction for multiple comparison, word-primed words also received marginally more delayed response than those double primed by pseudowords, $\beta = 0.03, SE = 0.01, t(90.50) = 2.42, p = .073$. However, neither single pseudoword-primed words nor double pseudoword-primed words produced any reliably longer response than unprimed words, $\beta = 0.01, SE = 0.01, t(41.39) = 1.22, p = .615$; $\beta = 0.02, SE = 0.01, t(60.54) = 1.41, p = .495$, and there was no significant difference between these two pseudoword-primed conditions either, $\beta = 0.002, SE = 0.01, t(143.32) = 0.20, p = .997$. Finally, post-hoc analyses of contrasts from the pseudoword target model did not show any significance after Tukey correction.

Accuracy. As shown in Figure 3.3B, responses to words have significantly higher error rate than pseudowords, $X^2(3) = 17.94, p < .001$. There was also a reliable interaction effect between lexicality and prime type, $X^2(2) = 15.13, p < .001$. Separate models for each lexicality type did not show any main effect of prime type for words, $X^2(3) = 3.93, p = .269$, but the effect for pseudowords was significant, $X^2(3) = 9.13, p = .028$. Since the pattern of these accuracy

priming effects were different from those in the MEG study, we conducted further post-hoc pairwise z tests on pseudoword targets, the results of which were Tukey corrected for multiple comparisons. These showed marginally higher accuracy caused by pseudoword priming compared to unprimed pseudowords, $\beta = 1.56$, $SE = 0.61$, $z = 2.56$, $p = .052$.

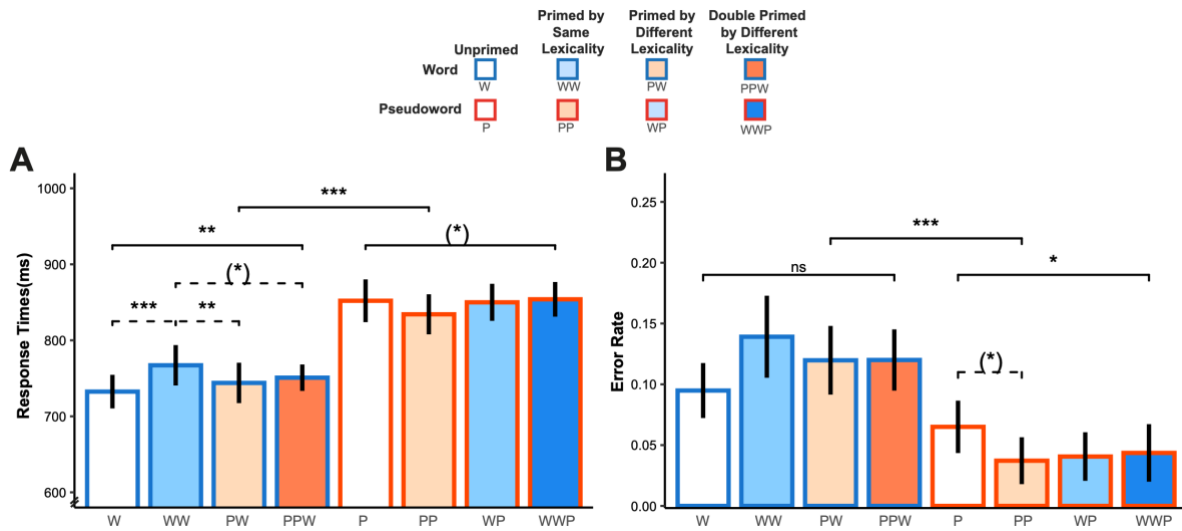


Figure 3.3. Response time results (**A**) and accuracy results (**B**) of the lexical decision task from Experiment 2. Bars are color-coded by lexicality and prime type on the x axis (words, blue frame; pseudowords, orange frame; unprimed, no fill; primed by same lexicality, consistent fill and frame colors; (double) primed by different lexicality, inconsistent fill and frame colors). Bars show the subject grand averages, error bars represent \pm within-subject CI, adjusted to remove between-subjects variance (Cousineau, 2005). Statistical significance is shown based on (generalised) linear mixed-effects regression in log-odds (for error rate) or log scale (for rt): (*) $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Statistical comparisons shown with solid lines indicate the lexicality by prime-type interaction and main effects of prime-type for each lexicality, while broken lines indicate the significance of pairwise comparisons.

3.2.3 Discussion

Overall, the results of Experiment 2 match the behavioural results from the MEG study, i.e. competitor priming affects words and pseudowords differently and the effect is only reliable for word-primed word. In addition, target items that were double primed by a different lexicality did not evoke reliably different response from those primed by a single item of different lexicality. Hence, it seems unlikely that the failure to observe significant pseudoword priming is due to this being too small an effect to detect following presentation of a single item.

However, the different word response times evoked by word and pseudoword priming may also be influenced by the lexical decision task itself, which may have implicitly cued participants to treat words and pseudowords differently. This is because participants need to identify the exact form of a single word in order to confirm its lexical status, but a deviation from all known words needs to be established to indicate a pseudoword (Norris & Kinoshita, 2008). In addition, lexical decision is also a decision made based on activity in semantic units (Plaut, 1997), namely participants' focus was mainly on lexical-semantic processing when performing this task. Therefore, we used a different task in Experiment 3, which is more “equal” for words and pseudowords while also directs participants' attention to sub-lexical processing.

3.3 Experiment 3

In order to address the question raised from Experiment 2, that is, whether the competitor priming effect is lexical-decision specific, we conducted Experiment 3 using the pause detection task (Davis et al., 2009; Dumay & Gaskell 2007; Gagnepain et al. 2012; Gaskell & Dumay, 2003; Mattys & Clark, 2002), which does not involve explicit linguistic judgement but is still sensitive to lexical activity. During the pause detection task, participants listen to the stimuli and then respond whether they have heard an unnatural pause in the stimuli. Mattys and

Clark (2002) first showed that response times during the pause detection task were affected by the number of activated lexical items at the location of the pause. For example, when pause was placed after a monosyllabic nonword segment, the response to the item was slower when the monosyllable matched more compatible lexical items. In addition, Gaskell and Dumay (2003) used the pause detection task to test the lexicalization of newly learned pseudowords. They found that slower pause detection response for existing words (e.g. *cathedral*) were only significant after long-term overnight consolidation of similar sounding pseudowords (e.g. *cathedruke*) but not after their immediate encoding. These results showed that only overnight consolidated pseudowords were added to the lexicon and confirmed that the pause detection task is sensitive to the change of overall amount of lexical activity (i.e. lexical entropy) before the pause location. Furthermore, Gagnepain et al. (2012) showed larger neural response after the pause location for words with lexicalized novel competitors, which indicates that this task is also sensitive to the size of prediction error.

Here, we used the pause detection task in Experiment 3 as a task that could direct participants attention to sub-lexical processing. However, since the task should reflect the change of lexical entropy before the pause location as well as the size of prediction error after pause, models such as TRACE, PC and DCM still predict a significant effect of competitor priming.

3.3.1 Methods

Participants

We recruited 65 paid participants from the same population as in Experiment 2 through Prolific Academic and all participants were tested online. This sample size was chosen based on a previous study using the pause detection task (Dumay & Gaskell, 2007). Five participants were

excluded from subsequent analyses due to error rate higher than two standard deviations above the mean, resulting in 61 participants in total.

Experimental Design and Materials

In Experiment 3, we adopted the same paradigm as Experiment 1 (i.e., without the double primed conditions) but used a pause detection task (Mattys & Clark, 2002) instead of the lexical decision task. Participants were asked to make judgements as to whether they heard an inserted pause in each item regardless of its lexicality via a keyboard response, with left index finger indicating pause absent and right index finger indicating pause present (Figure 3.4). In order not to over complicate the interpretation of results of interest, we kept all target words pause-absent and target pseudowords pause-present. Pause-present and -absent items were evenly distributed in prime words and pseudowords, with pause-absent prime words also served as the unprimed word condition and pause-present prime pseudowords also served as the unprimed pseudoword condition. Pause-present filler words and pause-absent filler pseudowords were added to balance the overall number of pause-present and -absent items in both lexicalities. Again, 480 trials were presented in 4 blocks, with each block of 120 trials lasting ~5 minutes.

The stimuli were the same as Experiment 2. In addition, for the purpose of the pause detection task, we inserted a pause of 200ms (Gaskell & Dumay, 2003) at the deviation point (DP) of each pause-present item.

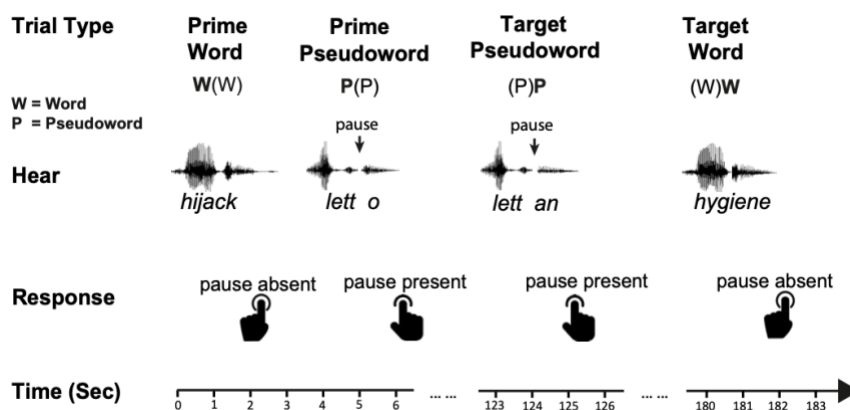


Figure 3.4. Experiment 3 task and design. A short pause of 200ms was inserted at the DP of each pause-present item.

Data analysis

Like in Experiment 2, all analyses were conducted using mixed-effect models to account for the random effect from participants and items. The RTs were measured from the DP, which was the same as the pause onset in items with pause (Mattys & Clark, 2002), and were log transformed to maximise normality. Exclusion criteria were the same as Experiment 2.

3.3.2 Results

Response Times. Figure 3.5A shows factorial analysis results of lexicality and prime type (unprimed, primed by same lexical status and primed by different lexical status). There was a significant main effect of lexicality, with pseudowords evoking longer response time than words, $X^2(3) = 98.86$, $p < .001$, the larger effect size (compared to previous experiments) is likely because all target pseudowords were pause present while words were pause absent. However, we did not find any reliable interaction between lexicality and prime type, $X^2(2) = 4.19$, $p = .123$. A one-way analysis of the words showed a significant prime type effect, $X^2(2) = 6.78$, $p = .034$, while there was no reliable prime effects for pseudowords, $X^2(2) = 0.04$, $p = .98$. We further examined the word pairwise effects, surprisingly, word-primed words did not evoke significantly longer RTs than unprimed words, $\beta = 0.01$, $SE = 0.01$, $t(89.72) = 0.85$, $p = .199$. Instead, pseudoword-primed words were recognised significantly more slowly than unprimed words, $\beta = 0.02$, $SE = 0.01$, $t(66.56) = 2.73$, $p = .008$. These results are inconsistent with predictions by the models and the competitor priming effect that we have found in Experiment 1 and 2 using the lexical decision task. In addition, the delayed word response

caused by pseudoword priming without extensive prior training or overnight consolidation is also contradictory to previous published studies (Dumay & Gaskell, 2007, Gaskell & Dumay, 2003).

Accuracy. Figure 3.5B shows that responses to words have significantly lower error rate than pseudowords, $X^2(3) = 72.38, p < .001$, which is understandably opposite to what we found using lexical decision task. There was no significant interaction effect between lexicality and prime type, $X^2(2) = 0.69, p = .707$, and we did not find any main effect of prime type for words, $X^2(2) = 0.32, p = .852$, or pseudowords, $X^2(2) = 0.73, p = .695$.

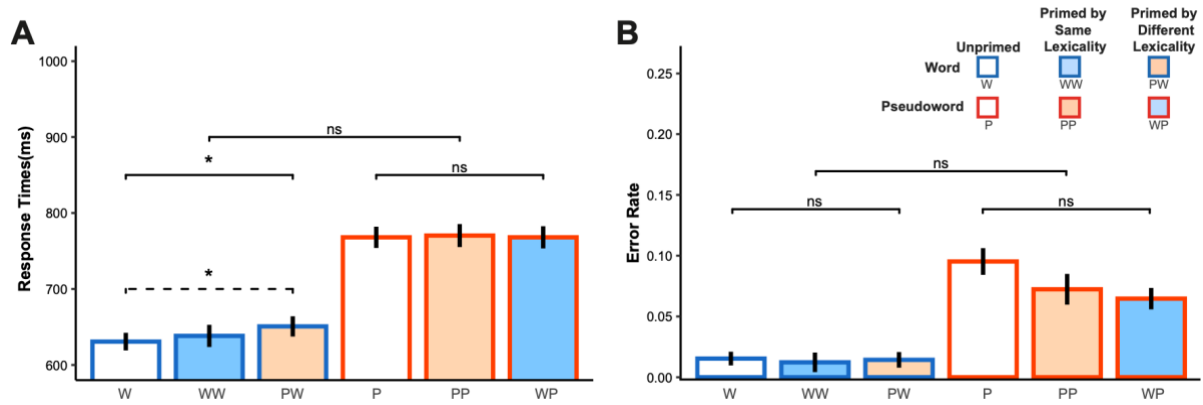


Figure 3.5. Response time results (A) and accuracy results (B) of the pause detection task from Experiment 3. Bars are color-coded by lexicality and prime type on the x axis (words, blue frame; pseudowords, orange frame; unprimed, no fill; primed by same lexicality, consistent fill and frame colors; primed by different lexicality, inconsistent fill and frame colors). Bars show the subject grand averages, error bars represent \pm within-subject CI, adjusted to remove between-subjects variance. Statistical significance is shown based on (generalised) linear mixed-effects regression: * $p < 0.05$. Statistical comparisons shown with solid lines indicate the lexicality by prime-type interaction and main effects of prime-type for each lexicality, while broken lines indicate the significance of pairwise comparisons.

3.3.3 Discussion

We conducted Experiment 3 with the goal of testing whether the word competitor priming effect was specific to the lexical decision task, during which words and pseudowords were likely processed differently. Our results showed that, when the task was changed to pause decision, word-primed words did not evoke reliably delayed response compared to unprimed words while pseudoword-primed words did.

We did not expect to observe the pseudoword priming effect since previous studies (Dumay & Gaskell, 2007; Gagnepain et al., 2012; Gaskell & Dumay, 2003) have demonstrated that for a pseudoword to be an effective ‘competitor’, it needs to be presented many more times during training and consolidated overnight or even over the course of several days. Hence, one-instance priming of a pseudoword did not seem to be sufficient to cause any delay in the target response immediately after priming. It may be that previous studies used a different task (phoneme detection) for training and that pause detection used during the perception of the prime item in this experiment somehow had different effects. Additionally, McMurray et al. (2016) also suggested that the setting of training (in isolation or with context) and the sensitivity of the task used for testing may affect whether newly learned items can be lexicalized immediately without overnight consolidation.

The non-significant word priming effect is also contrary to predictions made by spoken word recognition models such as TRACE, PC and DCM. One possible reason is that even though pause detection may reflect lexical activity indirectly, it needs stronger priming than just one prior presentation of the competitor word to evoke the competitor priming effect. As mentioned above, previous studies (Dumay & Gaskell, 2007; Gagnepain et al., 2012; Gaskell & Dumay, 2003) trained items many times – sometimes more than 10 times per item. However, this explanation makes the result contradictory to the significant delay caused by priming of

just one pseudoword item. Hence, it could also be that the experiment did not have large enough power to detect the competitor priming effect using the pause detection task.

Therefore, we conducted four more experiments (Experiment 4a, 4b, 5a, 5b) with increased number of participants based on power analyses results and increased number of presentations of each prime item so as to create stronger priming effects. We also dissociated the effect of lexical and sub-lexical processing more clearly by using separate prime and target phases and introduced pause in all prime items when using the pause detection task. With these changes in study design, we aimed to further examine in what way is sub-lexical processing different from lexical processing.

3.4 Experiment 4-5

In Experiment 2 and 3, we examined the competitor priming effect using lexical decision and pause detection tasks with prime and target items presented in an interleaved manner. However, the pause detection task is sensitive to lexical activity and prime words used in the word-primed word condition were always pause absent (i.e. the completeness of competitor prime words was not compromised) in Experiment 3, hence only the inhibitory effect of lexical processing (but not the facilitatory effect of phonological processing) was predicted by spoken word recognition models for both Experiment 2 and 3.

In order to better dissociate inhibitory lexical processing and facilitatory sub-lexical processing during priming and to test the influence of task-evoked decision (after word frequency has changed by competitor priming) at response to target items, we conducted a series of four more experiments (see Table 3.2 and Figure 3.6) with separate prime and target phases (necessary to avoid effects of task switching) and mixed the lexical decision and pause detection tasks in the two phases. Each prime item was presented four times to strengthen the priming effect and they were presented twice with pause and twice without pause during the

pause detection task. In this way, half of the presentations of the prime items were partially disrupted when the task was pause detection, hence there should be phonological segment processing without heavy influence from lexical processing of the complete item.

In experiment 4a and 4b, pause detection and lexical decision tasks were used respectively in each experiment in the prime phase, while the lexical decision task was used in the target phase for both experiments. In experiment 5a and 5b, prime phase also adopted the two different tasks respectively, while the task in the target phase was pause detection. The different tasks used in the prime phase in these experiments respectively tested lexical and sub-lexical processing during priming. The DCM (Gaskell & Marslen-Wilson, 1997) and PC (Davis & Sohoglu, 2020) predict an explicit facilitatory effect of phonological processing caused by the temporarily ambiguous pre-DP segments separated by pause during the pause detection task, and inhibitory effect of lexical competitor priming when the stimuli was presented without pause. TRACE (McClelland & Elman, 1986) also predicts the inhibitory lexical competitor priming effect, but the effect of facilitatory phonological priming may not directly affect spoken word recognition. In addition, the different tasks used in the target phase tested whether the lexical decision task was particularly more sensitive to the manipulation of word frequency (Balota & Chumbley, 1984) compared to the pause detection task. Note that Experiment 4b was a replication of Experiment 1 and 2 due to the use of lexical decision task in both phases. Although Experiment 5a also used the pause detection task in both phases, it was not an exact replication of Experiment 3 due to the existence of both pause-present and pause-absent presentations of each prime item.

For these experiments, we adopted a blended frequentist and Bayesian approach of data analysis. For results on word-primed words in all four experiments, we used the Bayesian framework to incorporate priors based on our previous study results of the competitor priming effect. This was also done for experiments 4a and 5a in which pause detection task was the

prime task, since even though phonological segments separated by pause might be processed through the sub-lexical pathway, pause-absent presentations for each item also existed and should be more sensitive to lexical processing. To avoid bias from priors, analyses using the frequentist approach were also conducted for all conditions of interest.

All four experiments were preregistered on OSF (Experiment 4a, [10.17605/OSF.IO/9453V](https://osf.io/9453V/); 4b, [10.17605/OSF.IO/JCUAR](https://osf.io/JCUAR/); 5a & b, [10.17605/OSF.IO/547GK](https://osf.io/547GK/)). Although it was mentioned in the preregistration documents that Experiment 4a and 4b would be analysed using the frequentist approach and Experiment 5a and 5b would be analysed using the Bayesian approach, we later decided that it was more reasonable to analyse all four experiments the same way, i.e. using the Bayesian approach only for word contrasts for which we had consistent *a priori* information (e.g. RT difference between word-primed word and unprimed word) while using the frequentist approach as a benchmark for all analyses in these experiments.

Table 3.2. Experiment tasks in 4a, 4b, 5a and 5b

Experiment	4a	4b	5a	5b
Prime Task	Pause Detection	Lexical Decision	Pause Detection	Lexical Decision
Target Task	Lexical Decision	Lexical Decision	Pause Detection	Pause Detection

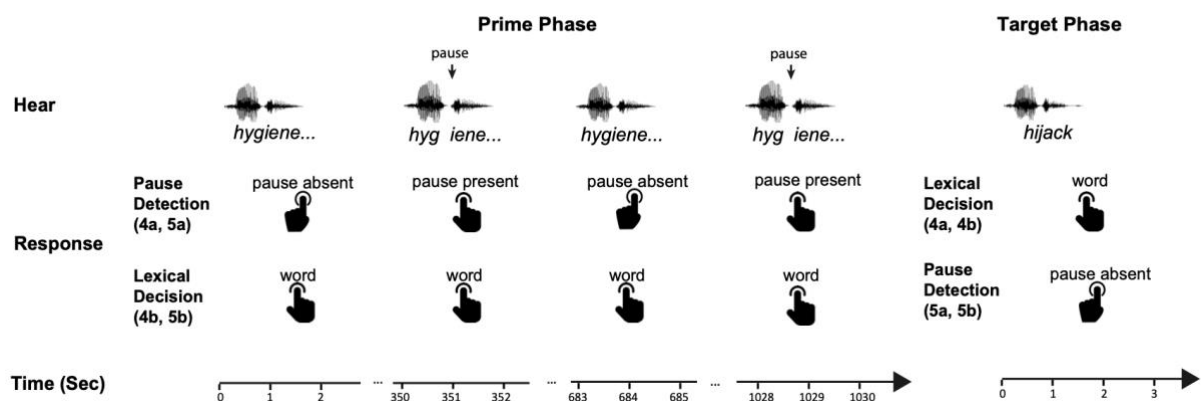


Figure 3.6. Paradigm of Experiment 4-5 with separate prime and target phases using different perception tasks.

3.4.1 Methods

Participants

A total of 409 paid participants from the same population pool as before were tested for these four experiments (88 on Experiment 4a, 86 on Experiment 4b, 118 on Experiment 5a, 117 on Experiment 5b). None of these participants had taken part in any of the previous experiments and they were all recruited through Prolific Academic and tested online. Participants with error rate two standard deviations above the mean were excluded from subsequent analyses, hence resulting in 81 on Experiment 4a, 77 on Experiment 4b, 110 on Experiment 5a and 110 on Experiment 5b.

We conducted power analyses to decide the sample size for Experiment 4 and 5. For Experiment 4a & b, we used the *simr* package (Green & MacLeod, 2016) implemented in R, which was specifically designed to calculate power for linear mixed models using simulation and to work with models fit using the *lme4* package (Bates et al. 2014). Based on the effect size of Experiment 2, we calculated the number of participants necessary to obtain 80% power to detect a slope of 0.05 (log-scale) at a 0.05 significance level. For Experiment 5a & b, the sample size was calculated using the sequential design of the Bayes Factor Design Analysis (BFDA, Schoenbrodt & Wagenmakers, 2018; Stefan et al., 2019) since we adopted a blended Bayesian approach for the analyses of these experiments (see *Data Analysis* section). Using a neutral design prior that follows a central Cauchy distribution (with a location at 0 and a scale parameter of $\sqrt{2}/2$) and based on an estimated effect size of the pseudoword-primed word > unprimed word contrast in Experiment 3, we simulated 10000 hypothetical studies using the

BFDA package implemented in R (Schoenbrodt & Stefan, 2018), with the Bayes factor boundary set to 6 (H1) and 1/6 (H0). As shown in Figure 3.7, our simulation results indicated that a sample size of ~110 is needed for ~80% of the simulated studies to hit the H1 boundary with BF 6 as the threshold, and ~75 is needed to hit the corresponding H0 boundary. In order to be able to collect high quality data from 110 participants, we set the maximum number of sample size to 120 participants for each study and, based on the sample size of Exp4a & b, we tested a minimum number of 80 participants, and then kept testing until reaching either of the stopping boundary or the maximum number of sample size.

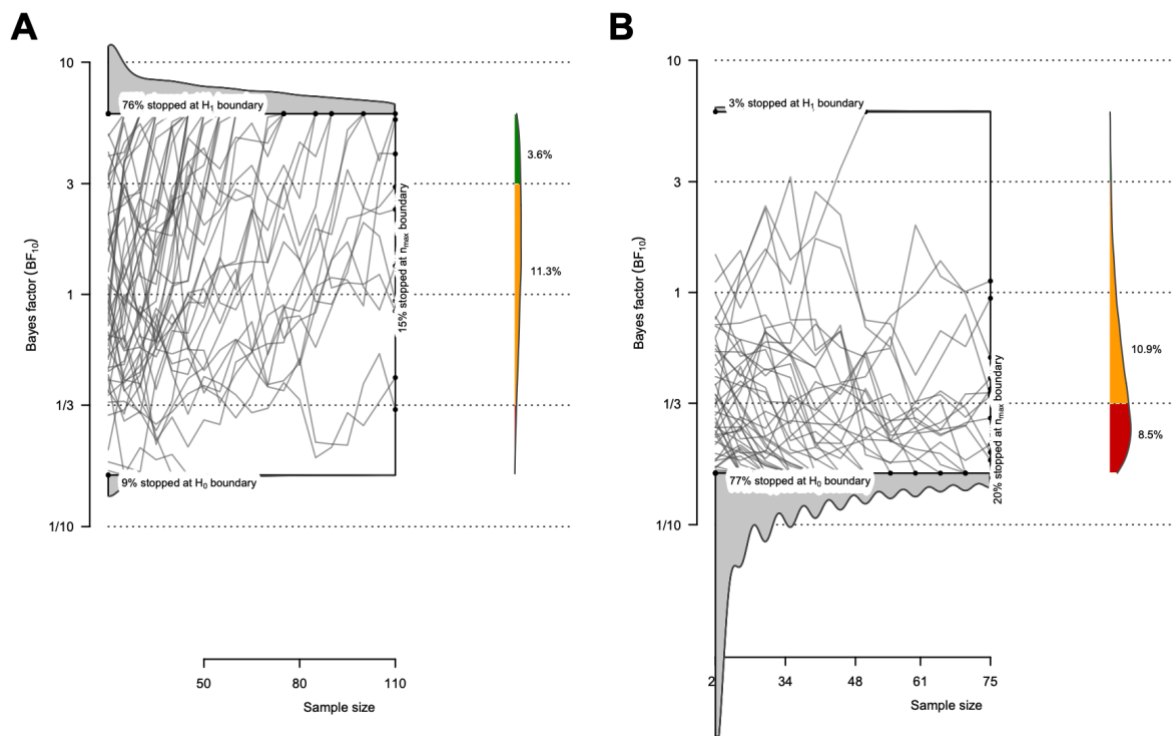


Figure 3.7. Sample size calculated using the Bayes Factor Design Analysis (BFDA) based on 10000 simulations. The Bayes factor boundaries are set to 6 (H1) and 1/6 (H0). A. Sample size based on simulations for H1. B. Sample size based on simulations for H0.

Experimental Design and Materials

Experiment 4 and 5 did not have prime and target items presented in an interleaved manner as in previous experiments, instead, these experiments have separate prime and target phases (see Figure 3.6). This way, lexical and sub-lexical processing could be studied using different tasks within the same experiment and the effect of different kinds of priming could also be compared across experiments. As shown in Table 3.2, Experiment 4a and b adopted the pause detection task and the lexical decision task respectively in the prime phase, while both used the lexical decision task in the target phase. Experiment 5a and b also used different tasks in their prime phase, but both used pause detection task in the target phase.

In the prime phase, the stimuli (the same as in Experiment 1-3) were presented in 4 blocks of 108 trials, with 54 words and 54 pseudowords in each block. Two filler items were added at the beginning of each block to help participants get ready. Each spoken item was repeatedly presented across the four blocks, hence 4 repetitions per item (i.e. one presentation of each item in each block). If pause detection task was used in the prime phase, then each item was presented with and without pause twice respectively.

In the target phase, unprimed control items and target items sharing the same initial segments with prime were presented, among which each prime type condition (word-primed word, word-primed pseudoword, pseudoword-primed word, pseudoword-primed pseudoword, unprimed word, unprimed pseudoword) contains 27 unique target items and each item was presented only once. If pause detection task was used in the target phase, then like in Experiment 3, target words were pause absent and target pseudowords were pause present. In Experiment 4a and 4b, 27 word fillers and 27 pseudoword fillers were added so that half of the items in the target phase had the same initial sounds as prime, while the other half had different beginnings. In Experiment 5a and 5b, 81 word fillers with inserted pause and 81 pseudoword fillers without pause were added so that both words and pseudowords have the same number of items with and without pause.

Note that the lag between the last prime block and the target phase was longer than the lag between prime and target items presented in an interleaved manner in previous experiments. In Experiment 4a and 4b, the lag range was between ~20 and ~310 trials, which was between ~1min to ~10min. In Experiment 5a and 5b, due to the added number of fillers, the lag range was between ~20 and ~410 trials, approximately from ~1min to ~14min.

Data analysis

In Experiment 4 and 5, we adopted a blended frequentist and Bayesian approach (Bickel, 2015). While the frequentist framework was the same as in Experiment 2 and 3, the Bayesian analysis framework was implemented in the probabilistic programming language *Stan* (Carpenter et al., 2017) using the model wrapper package *brms* in R (Bürkner, 2017). Our motivation for using this blended approach is that the Bayesian method allows us to incorporate insights from our previous studies using priors (e.g. competitor priming effect for words), while the frequentist method serves as a benchmark when there is no reliable prior knowledge (e.g. response accuracy and pseudowords RTs) and when bias from the prior needs to be avoided. Hence, we think that it makes sense to report posterior results from Bayesian models when the prior is known (either based on previous literature or experiment results) and to report inference from frequentist models for all conditions of interest at the same time for when there is little knowledge of the prior and to make sure that the results were not biased by unreliable priors that are too narrow or too big.

We assessed convergence of Bayesian models by verifying that there were no divergent transitions and that the *Rhat* (between- to within-chain variances) was close to one. As before, we fit linear and logistic mixed effect models on transformed RTs and error rates respectively. Mean estimates (β), 95% quantile-based Bayesian credible intervals (CrI, an interval that contains the true value with 95% probability) and Bayes factors (a measurement that quantifies

the evidence in favour of one hypothesis over another, e.g. a null hypothesis relative to an alternative hypothesis) were reported from Bayesian models.

Apart from analysing the main effect of lexicality and the interaction between lexicality and prime type as in previous experiments, our design of Experiment 4a & b and 5a & b also allowed us to combine the data from these 4 experiments respectively to compare the effect of different prime tasks (lexical decision vs pause detection) while controlling for the target task and compare the effects of different target tasks while controlling for the prime task.

To evaluate the main effects, we compared the full model to nested reduced models using the likelihood-ratio tests under the frequentist approach. The effect of planned pairwise contrasts were evaluated using Bayes factors calculated through the Savage-Dickey density ratio method (Wagenmakers et al., 2010). In order to make sure that Bayes factors were not biased by informative priors, we also checked these results against the frequentist tests.

On top of the exclusion criteria used for previous experiments, in Experiment 4 and 5, prime items with an error rate of more than 50% (more than two incorrect responses out of four) were excluded, so were their corresponding target items.

3.4.2 Results of Experiment 4a (prime: pause detection; target: lexical decision)

Response Times. RT results from the target phase (lexical decision task) of Experiment 4a are shown in Figure 3.8A. The factorial analysis of lexicality and prime type was conducted using the frequentist model. As before, pseudowords evoked significantly longer response time than words, $X^2(3) = 31.08$, $p < .001$, but the interaction between lexicality and prime type was not reliable, $X^2(2) = 1.11$, $p = .574$. In addition, there was no main effect of prime type for either words, $X^2(2) = 2.28$, $p = .320$, or pseudowords, $X^2(2) = 0.19$, $p = .908$.

With the hypothesis that lexical competitor priming effect can be evoked on word-primed word conditions regardless of which task being used in either prime or target phase, we fitted a Bayesian model on these contrasts to incorporate our prior knowledge using informative priors. As shown in Figure 3.9A, we used a normal distribution with 0.045 as the mean and 0.02 as the standard deviation ($Normal(0.045, 0.02)$) as the prior for word-primed word > unprimed word condition. This prior distribution is on log scale since the RTs are log-transformed. The mean (i.e. the location of the distribution) indicated an effect size of ~30ms difference between the RTs evoked by word-primed words and unprimed words (the reference condition) based on our Experiment 1 and 2. The standard deviation, also the scale of the distribution, limited the range of the effect size to ~3ms - ~59ms, which is informative on the direction of the effect while not overly restrictive given the surprising results we have found in Experiment 3. Figure 3.9B shows that the prior for pseudoword-primed word > unprimed word condition was neutral, $Normal(0, 0.1)$, as we did not have a clear prediction on pseudoword priming effects.

With these priors, the posterior distribution showed that word-primed word evoked longer RTs than unprimed word, $\beta = 0.02$, 95% CrI = [0.001, 0.03]. However, since the evidence for 0 is still stronger in the posterior distribution than in the prior distribution, Bayes factor analysis using the Savage-Dickey density ratio method showed moderate support for the null hypothesis ($BF_{01} = 3.5$), which is consistent with the marginal result under the frequentist framework, $\beta = 0.01$, $SE = 0.01$, $t(205.4) = 1.39$, $p = .083$. Pseudoword-primed word did not show a clear effect in delaying the response, $\beta = 0.002$, 95% CrI = [-0.02, 0.02], and the Bayes factor analysis showed strong support for the null ($BF_{01} = 11.6$).

These results did not seem to indicate any clear effect caused by word competitor priming and there was no reliable effect of pseudoword-priming on targets. The use of pause detection task in the prime phase, with both pause absent and pause present presentations of

prime items, should trigger both lexical- and phonological-level processing. The use of lexical decision task in the target phase should be particularly sensitive to the manipulation of word frequency (Balota & Chumbley, 1984). The marginal delay of recognition on word-primed word seemed to show that both inhibitory effect of lexical competitor priming and facilitatory effect of phonological priming played a role during this process. It may also be possible that the lag between prime and target was too long to evoke any pronounced effect. Experiment 4b using the same lexical decision task in both prime and target phase helped us further address these questions.

Accuracy. Figure 3.8B shows that responses to words have significantly higher error rate than pseudowords, $X^2(3) = 28.88, p < .001$. There was no significant interaction effect between lexicality and prime type, $X^2(2) = 1.42, p = .490$. Main effect of prime type was found for word conditions, $X^2(2) = 6.88, p = .032$, specifically unprimed words had significantly higher accuracy than pseudoword-primed words after Tukey correction, $\beta = 0.29, SE = 0.12, z = 2.49, p = .034$. We did not find any significant main effect of prime type for pseudowords, $X^2(2) = 0.11, p = .945$.

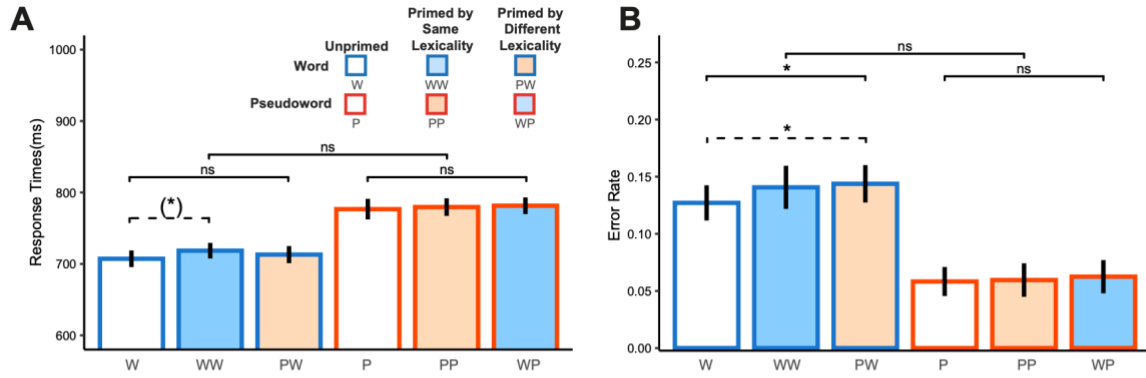
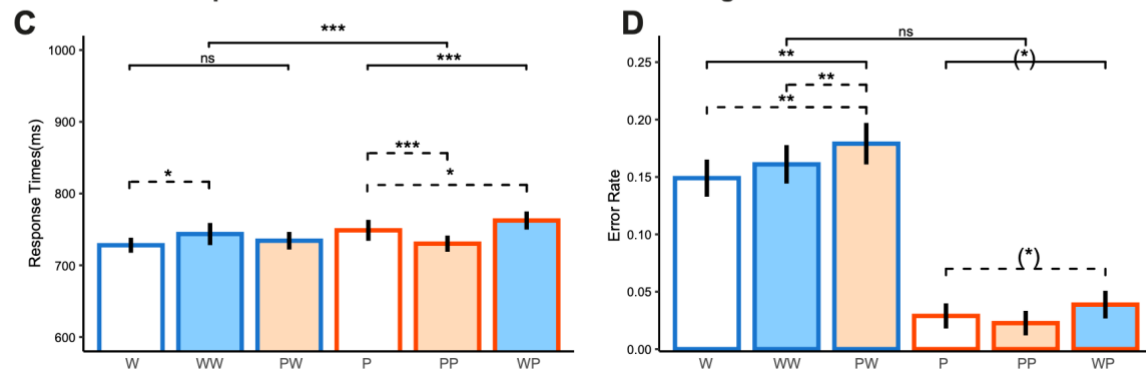
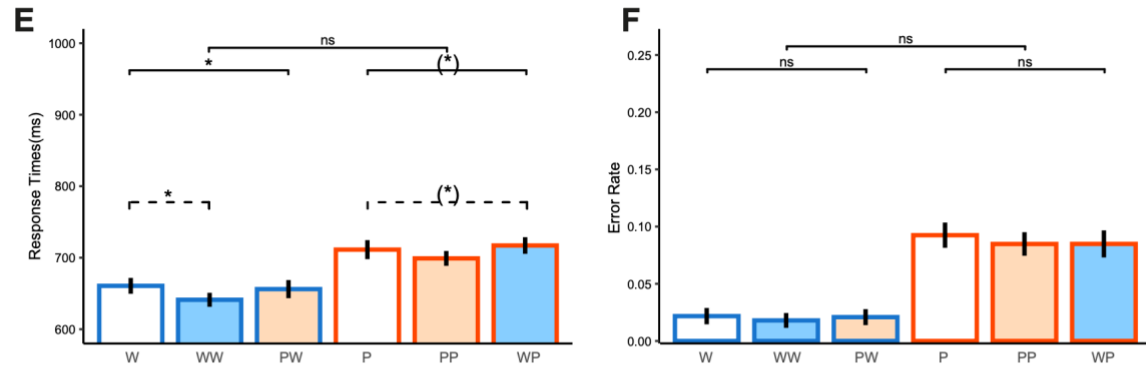
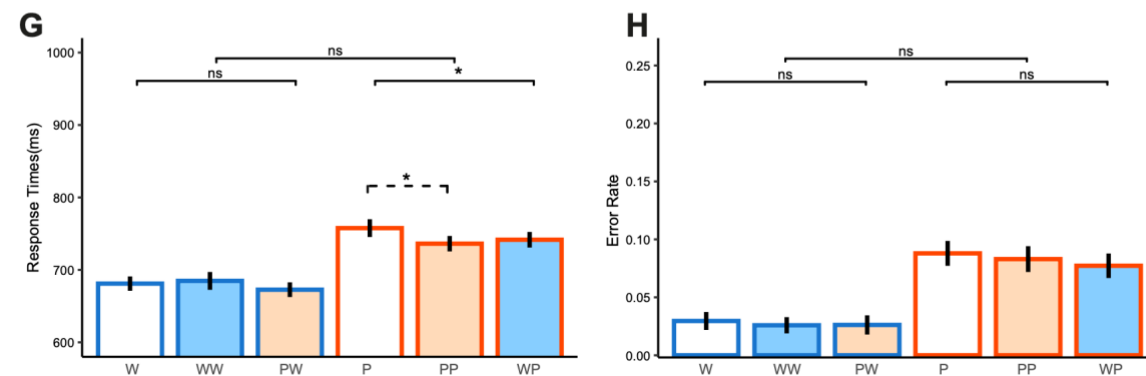
Experiment 4a Prime: Pause Detection - Target: Lexical Decision**Experiment 4b Prime: Lexical Decision - Target: Lexical Decision****Experiment 5a Prime: Pause Detection - Target: Pause Detection****Experiment 5b Prime: Lexical Decision - Target: Pause Detection**

Figure 3.8. Response time results (*A, C, E, G*) and accuracy results (*B, D, F, H*) of the target phase in Experiment 4-5. Bars are color-coded by lexicality and prime type on the x axis (words, blue frame; pseudowords, orange frame; unprimed, no fill; primed by same lexicality, consistent fill and frame colors; primed by different lexicality, inconsistent fill and frame colors). Bars show the subject grand averages, error bars represent \pm within-subject CI, adjusted to remove between-subjects variance. Statistical significance is shown for error rate results based on generalised linear mixed-effects regression: ** $p < 0.01$, * $p < 0.05$, (*) $p < 0.1$. Statistical comparisons shown with solid lines indicate the lexicality by prime-type interaction and main effects of prime-type for each lexicality, while broken lines indicate the significance of pairwise comparisons. Statistical significance is not shown for RT results due to the blended analyses approach.

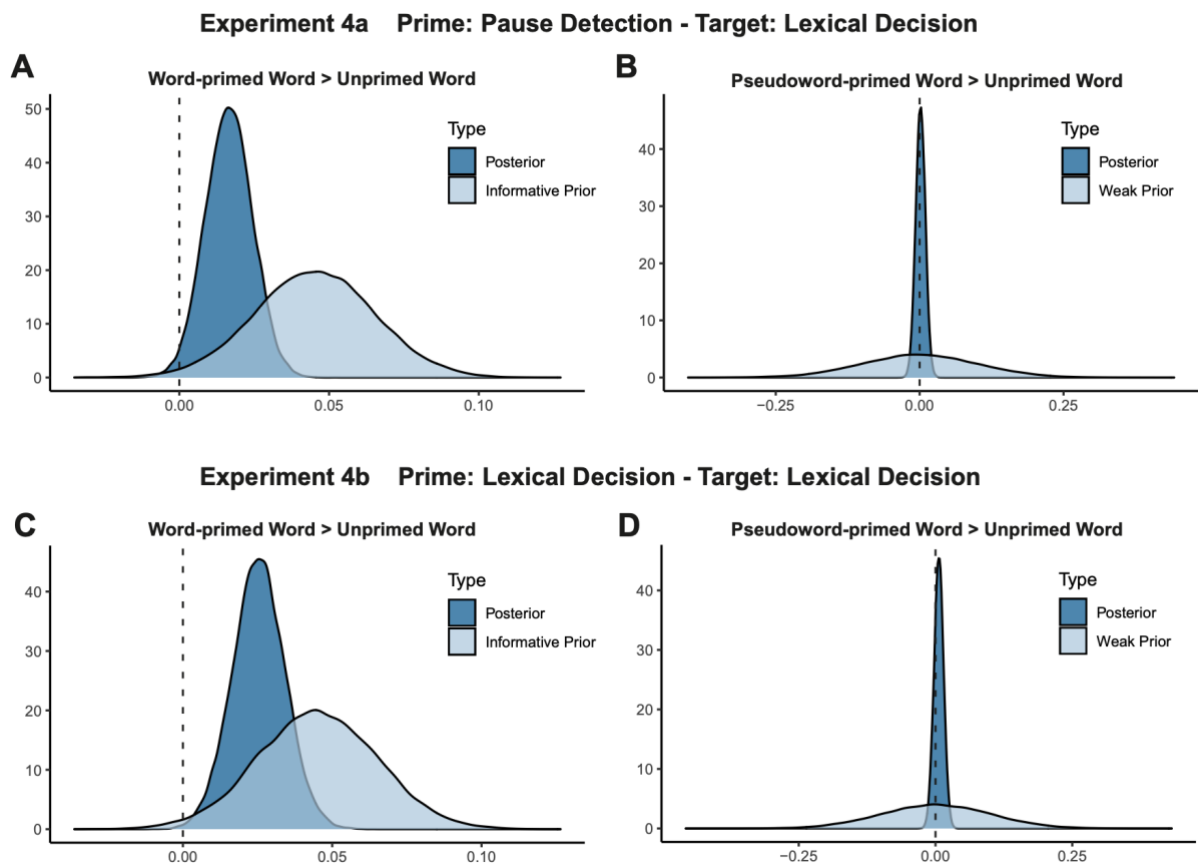


Figure 3.9. Prior and posterior distribution for word-primed word > unprimed word RTs and pseudoword-primed word > unprimed word RTs in Experiment 4. X-axis shows log RT differences, y-axis shows the density of the distributions. The prior distribution shows the initial probability of the contrast difference before seeing the data, while the posterior distribution shows the probability of the contrast difference given the data. **A, C.** The word-primed word > unprimed word contrast in Experiment 4a & b had the same informative prior, $Normal(0.045, 0.02)$, which was based on results from Experiment 1 and 2. **B, D.** The pseudoword-primed word > unprimed word contrasts in these two experiments had a neutral prior, $Normal(0, 0.1)$.

3.4.3 Results of Experiment 4b (prime: lexical decision; target: lexical decision)

Response Times. RT results from the target phase (lexical decision task) of Experiment 4b is shown in Figure 3.8C. The factorial analysis of lexicality and prime type showed that pseudowords evoked significantly longer response time than words, $X^2(3) = 33.30, p < .001$. There was also a significant interaction between lexicality and prime type, $X^2(2) = 45.07, p < .001$. The main effect of prime type was not significant for words, $X^2(2) = 4.30, p = .117$, but surprisingly it was significant for pseudowords, $X^2(2) = 28.83, p < .001$. More specifically, after Tukey correction, pseudoword-primed pseudowords evoked significantly quicker response than unprimed pseudowords, $\beta = -0.03, SE = 0.01, t(151.27) = -3.69, p < .001$, while word-primed pseudowords evoked significantly delayed responses than unprimed ones, $\beta = 0.02, SE = 0.01, t(151.55) = 2.49, p = .035$. These could be practice effects caused by associative learning between stimuli and responses. Since each item with unique initial segments corresponded to either word or pseudoword response four times during the prime phase, it is possible that participants already associated certain initial segments with certain

lexical decision, hence priming by the same lexicality facilitated lexical decision on the target item while priming by a different lexicality caused delay in its recognition.

We then conducted Bayesian analyses on the RT results of word contrasts. As shown in Figure 3.9 C & D, we used the same informative prior, $Normal(0.045, 0.02)$, as in Experiment 4a for word-primed word > unprimed word condition, and the same neutral prior for pseudoword-primed word > unprimed word condition, $Normal(0, 0.1)$. We found that word-primed word evoked longer RTs than unprimed word, $\beta = 0.03$, 95% CrI = [0.01, 0.04]. The Bayes factor was also weakly in favour of the alternative hypothesis ($BF_{10} = 2.2$), which was consistent with the results under the frequentist framework, $\beta = 0.02$, $SE = 0.01$, $t(79.06) = 2.02$, $p = .023$. Pseudoword-primed word did not show a clear effect in delaying the response, $\beta = 0.01$, 95% CrI = [-0.01, 0.02], and the Bayes factor analysis showed strong support for the null ($BF_{01} = 8.9$).

The word competitor priming effect was reliable although the effect size was smaller than those in Experiment 1 and 2. One possible reason is that the same kind of practice effect as shown for pseudoword targets might have dampened the effect size. Another explanation is that the longer lag between prime and target items reduced the strength of the effect.

Accuracy. Figure 3.8D shows significantly greater error rates for word responses than for pseudoword responses, $X^2(3) = 65.14$, $p < .001$. There was no significant interaction effect between lexicality and prime type, $X^2(2) = 4.27$, $p = .119$. The main effect of prime type was found for word conditions, $X^2(2) = 13.49$, $p = .001$, specifically unprimed words had significantly higher accuracy than pseudoword-primed words, $\beta = 0.57$, $SE = 0.16$, $z = 3.54$, $p = .001$, and word-primed words, $\beta = 0.57$, $SE = 0.18$, $z = 3.23$, $p = .004$, both survived Tukey correction. This pattern is similar to that found in Experiment 4a. In addition, there was a marginal main effect of prime type for pseudowords, $X^2(2) = 5.84$, $p = .054$, in which word-

primed pseudoword evoked marginally higher error rate than unprimed pseudoword, $\beta = 1.16$, $SE = 0.51$, $z = 2.29$, $p = .058$.

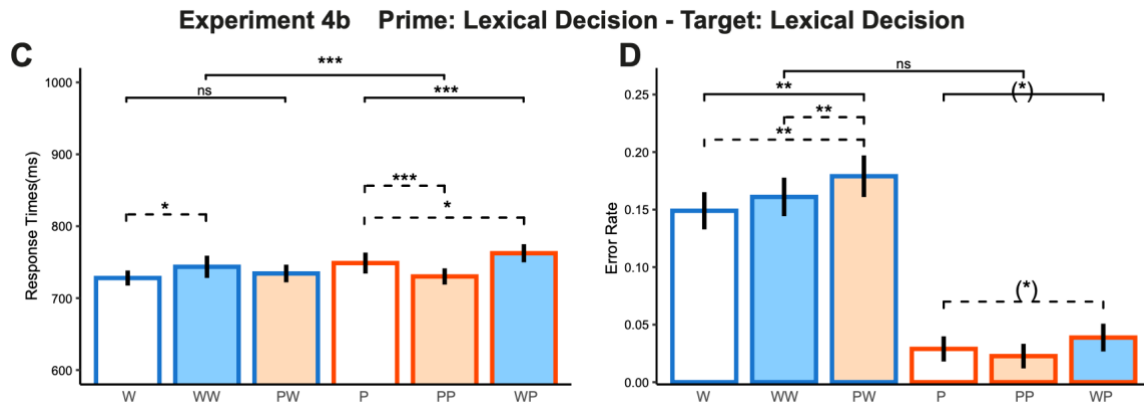


Figure 3.8C & D

3.4.4 Results of Experiment 5a (prime: pause detection; target: pause detection)

Response Times. RT results from the target phase (pause detection task) of Experiment 5a are shown in Figure 3.8E. Like before, pseudowords evoked significantly longer response time than words, $X^2(3) = 37.73$, $p < .001$, although the interaction between lexicality and prime type was not significant, $X^2(2) = 1.15$, $p = .562$. The main effect of prime type was significant for words, $X^2(2) = 6.70$, $p = .035$, and marginal for pseudowords, $X^2(2) = 4.95$, $p = .084$, with word-primed pseudowords evoking marginally delayed response than pseudoword-primed pseudowords after Tukey correction, $\beta = 0.02$, $SE = 0.01$, $t(140.9) = 2.25$, $p = .064$.

For the Bayesian analyses on word target RT results, we used an informative prior, $Normal(0.017, 0.008)$, based on the results of word-primed word > unprimed word condition from Experiment 4a (since both experiments used pause detection as the prime task), as shown in Figure 3.10A. The effect of word priming compared to the unprimed condition was unclear, $\beta = 0.001$, 95% CrI = $[-0.01, 0.01]$, and since a large portion of posterior samples fell outside

of the prior distribution, this informative prior might have been too restrictive. Hence we switched to a weaker prior using a larger standard deviation (wider scale), $Normal(0.017, 0.1)$, as shown in Figure 3.10B, to prevent the prior from biasing the posterior results too much. The results indicate that word-primed words evoked faster response than unprimed words, $\beta = -0.02$, 95% CrI = [-0.04, -0.003]. The Bayes factor also weakly supported the alternative hypothesis ($BF_{10} = 1.4$). This result matched with the reliable effect under the frequentist framework, $\beta = -0.02$, $SE = 0.01$, $t(111.18) = -2.26$, $p = .026$. For the pseudoword-primed word > unprimed word condition, we used the same neutral prior, $Normal(0, 0.1)$, as shown in Figure 3.10C. Pseudoword priming did not show a clear effect although it is also in the negative direction, $\beta = -0.01$, 95% CrI = [-0.03, 0.02], and the Bayes factor analysis showed strong support for the null hypothesis ($BF_{01} = 8.3$).

This experiment used the pause detection task in both prime and target phases and showed shortened RTs evoked by word primed word. The finding from Experiment 3 that pseudoword priming caused delayed word response was not replicated. The faster response caused by word priming here was not due to associative learning between stimuli and response either, since each prime word was presented twice with pause and twice without pause. And it is exactly because prime items were presented in both ways, together with the use of the pause detection task, that it was possible for these stimuli to be processed at both lexical level and sub-lexical level during the prime phase. Additionally, even though the pause detection task was found to be sensitive to the change of lexical entropy (Gaskell & Dumay, 2003) and prediction error (Gagnepain et al., 2012), using the pause detection task in the target phase still evoked different effects from using the lexical decision task. Taken together, the shorter response times evoked by word-primed word seemed to reflect the facilitatory effect from sub-lexical processing of the pre-DP segment. However, it is also unclear why similar phonological facilitatory effect was not shown for pseudoword-primed words. It may be that the unfamiliar

post-DP segment in pseudowords reduced the activation of words compatible with the pre-DP segment during the pause detection task.

Accuracy. Opposite to Experiment 4a & b, Figure 3.8F shows that pseudowords evoked significantly higher error rate than words, $X^2(3) = 42.77$, $p < .001$. There was no significant interaction effect between lexicality and prime type, $X^2(2) = 1.35$, $p = .509$. Main effect of prime type was not significant for either word targets, $X^2(2) = 1.45$, $p = .485$, or pseudoword targets, $X^2(2) = 3.05$, $p = .218$.

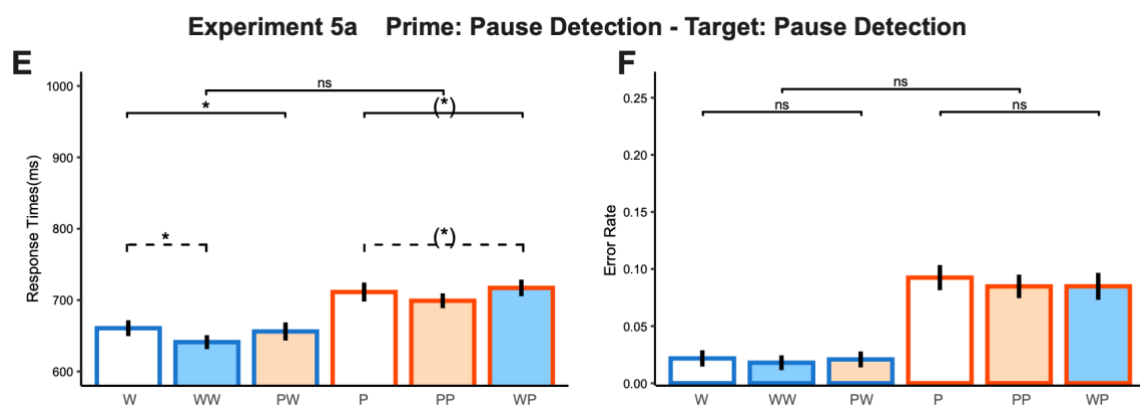


Figure 3.8E & F

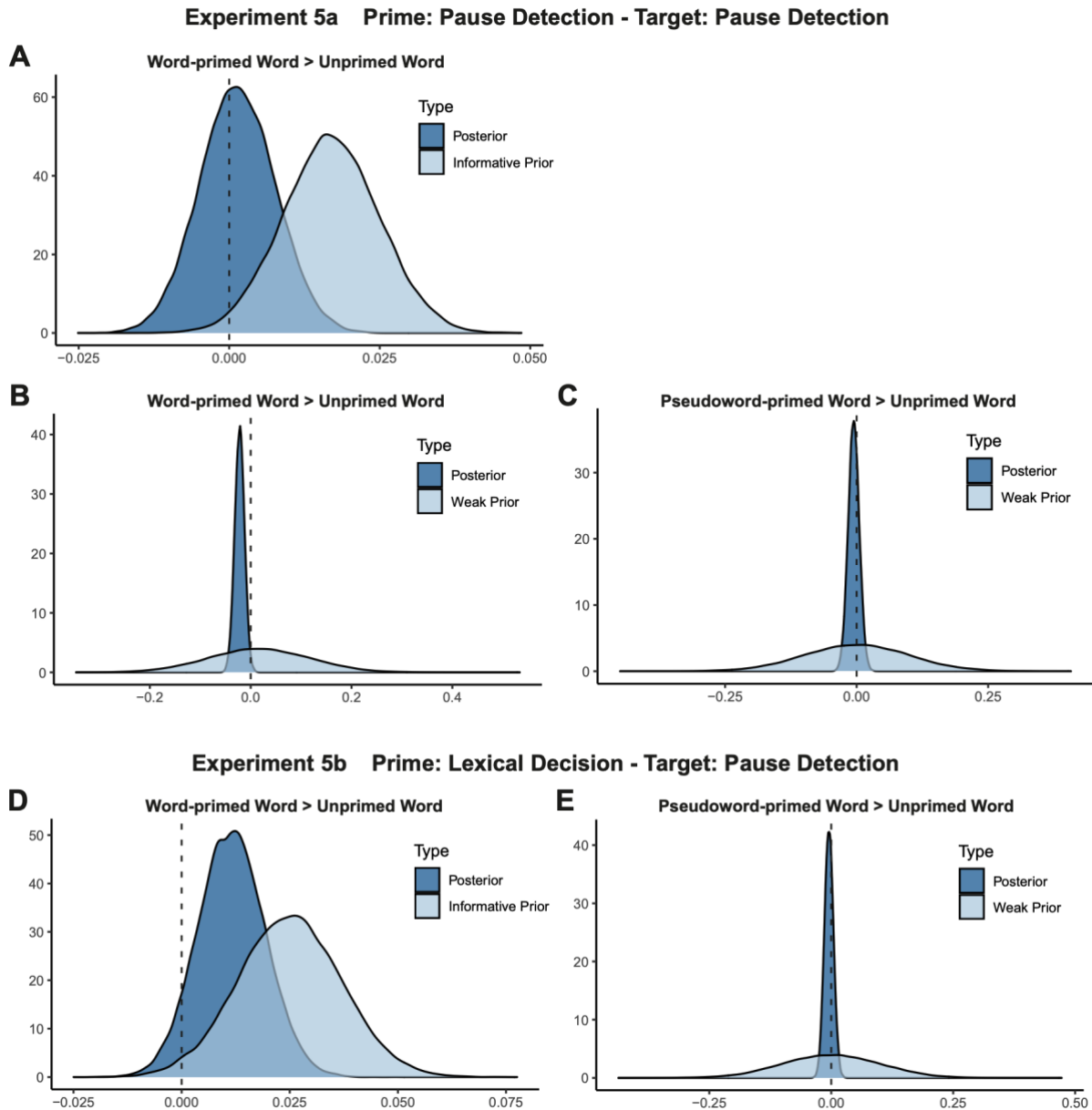


Figure 3.10. Prior and posterior distribution for word-primed word > unprimed word RTs and pseudoword-primed word > unprimed word RTs in Experiment 5. X-axis shows log RT differences, y-axis shows the density of the distributions. The prior distribution shows the initial probability of the contrast difference before seeing the data, while the posterior distribution shows the probability of the contrast difference given the data. **A.** An informative prior, $Normal(0.017, 0.008)$, was applied to the word-primed word > unprimed word contrast in Experiment 5a, based on results from Experiment 4a. **B.** A weaker prior, $Normal(0.017, 0.1)$, was applied to the word-primed word > unprimed word contrast in Experiment 5a as the

informative prior appears to be too restrictive. **D.** An informative prior, $Normal(0.025, 0.012)$, was applied to the word-primed word > unprimed word contrast in Experiment 5b, based on results from Experiment 4b. **C, E.** A neutral prior, $Normal(0, 0.1)$, was applied to the pseudoword-primed word > unprimed word contrast in both Experiment 5a and 5b.

3.4.5 Results of Experiment 5b (prime: lexical decision; target: pause detection)

Response Times. RT results from the target phase (pause detection task) of Experiment 5b is shown in Figure 3.8G. Pseudowords evoked significantly longer response time than words, $X^2(3) = 54.47, p < .001$, but the interaction between lexicality and prime type was not reliable, $X^2(2) = 4.08, p = .13$. The main effect of prime type was not significant for words, $X^2(2) = 1.89, p = .388$, but it was significant for pseudowords, $X^2(2) = 7.48, p = .024$, with pseudoword-primed pseudowords evoking significantly faster response than unprimed pseudowords after Tukey correction, $\beta = -0.02, SE = 0.01, t(90.80) = -2.69, p = .020$. Although this effect is reminiscent of the potential practice effect shown for pseudoword targets in Experiment 4b, it is worth noting that in Experiment 5b, different tasks were used in the prime phase and the target phase, hence even if an association was created between the stimuli and the lexicality response during the prime phase, it should not affect the pause detection task in the target phase. I will discuss more on what this pseudoword priming effect indicates in the General Discussion.

For the Bayesian analyses on word contrasts, we used an informative prior, $Normal(0.025, 0.012)$, based on the results of word-primed word > unprimed word condition from Experiment 4b (since both experiments used the lexical decision task in the prime phase), as shown in Figure 3.10D. Here, word priming effect did not seem to be reliable, though it was in the positive direction, $\beta = 0.01, 95\% CrI = [-0.004, 0.027]$. The Bayes factor analysis showed

moderate support for the null hypothesis ($BF_{01} = 4.2$). These results were in line with the frequentist test, $\beta = 0.002$, $SE = 0.01$, $t(132.68) = 0.20$, $p = .842$. For the pseudoword-primed word > unprimed word condition, we applied the same neutral prior, $Normal(0, 0.1)$, as shown in Figure 3.10E. Pseudoword-primed word did not show a clear effect compared to the unprimed condition although it is in the negative direction, $\beta = -0.01$, 95% CrI = $[-0.022, 0.013]$, and the Bayes factor analysis showed strong support for the null ($BF_{01} = 9.82$).

These results showed that even when the lexical decision task was used in the prime phase, using the pause detection task in the target phase would not lead to a significant competitor priming effect. One possible reason is that, as Balota and Chumbley (1984) suggested, the lexical decision task is especially sensitive to the manipulation of word frequency compared to other tasks. However, we could not rule out the possibility that the long lag (even longer than in Experiment 4a and 4b) between prime and target items made it impossible to evoke reliable lexical competitor priming effect.

Accuracy. Similar to the pattern shown in Experiment 5a, Figure 3.8H indicates that pseudowords evoked significantly higher error rate than words, $X^2(3) = 37.75$, $p < .001$. There was no significant interaction effect between lexicality and prime type, $X^2(2) = 1.14$, $p = .564$. Main effect of prime type was not significant for either word targets, $X^2(2) = 3.25$, $p = .197$, or pseudoword targets, $X^2(2) = 2.23$, $p = .327$.

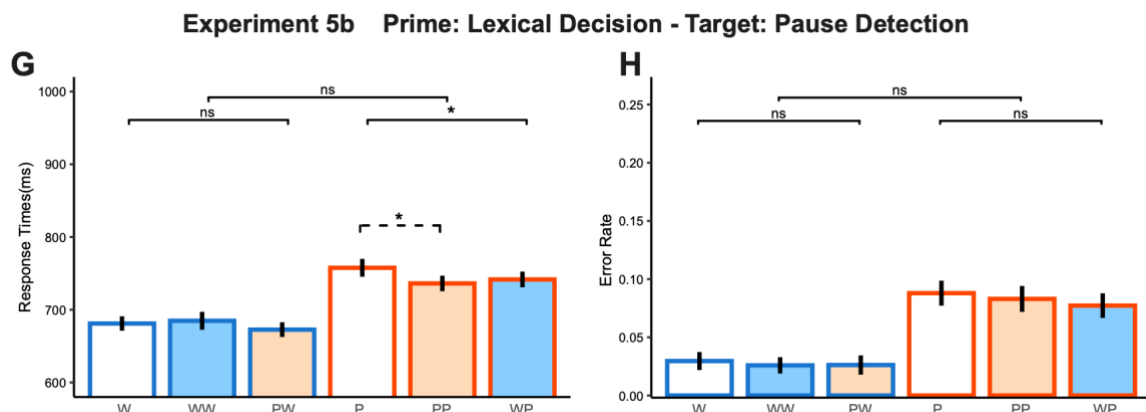


Figure 3.8G & H

3.4.6 Comparison of Word Targets RT results from Experiment 4 & 5

As mentioned in the Method section (3.4.1), our design of Experiment 4 and 5 also allowed us to compare different prime task effects while controlling for the target task and vice versa. This is especially helpful for dissociating the effect of lexical and sub-lexical processing that we led the participants to focus on during the prime and target phase. A comparison of RT effect size on word-primed word > unprimed word and pseudoword-primed word > unprimed word under the Bayesian framework from all four experiments is shown in Figure 3.11.

We then combined RT data from word targets in all four experiments and included a prime task factor (lexical decision, pause detection) and a target task factor (lexical decision, pause detection) on top of the prime type factor in our linear mixed-effect models. In addition, we also conducted analyses on combined data from paired experiments (Experiment 4a and 4b, 4a and 5a, 5a and 5b, 4b and 5b), with either prime task or target task as an additional factor. Since these analyses involve the new task factor, they were conducted under the frequentist framework.

Our word model with all data together and controlling for the prime and target tasks showed that word priming significantly delayed word recognition, $\beta = 0.02$, $SE = 0.01$, $t(212.1) = 2.21$, $p = .029$. In addition, pause detection as the target task led to significantly faster response than when lexical decision was used as the target task, $\beta = -0.08$, $SE = 0.03$, $t(468.5) = -2.98$, $p = .003$.

The word model with combined data from 4a and 4b (controlling for the target task) did not show any reliable difference in RTs caused by the different prime tasks, $\beta = 0.02$, $SE = 0.02$, $t(160.8) = 0.74$, $p = .459$, hence we removed the prime task factor from this model and

ran a model with prime type as the only factor. This model also showed a significant delay in word recognition caused by word priming, $\beta = 0.02$, $SE = 0.01$, $t(134.7) = 2.28$, $p = .024$.

The model with data from Experiment 4b and 5b (controlling for the prime task) showed a significant effect of target task, $\beta = -0.08$, $SE = 0.03$, $t(212.4) = -2.83$, $p = .005$, consistent with the results from the model with all data. In addition, it also indicated that word-primed words evoked significantly longer RTs than unprimed words, $\beta = 0.02$, $SE = 0.01$, $t(121.5) = 2.11$, $p = .037$. These results indicated that Experiment 4a, 4b and 5b shared the same pattern in terms of the competitor priming effect.

On the other hand, the model with combined data from Experiment 4a and 5a (controlling for the prime task) showed a significant effect of target task, $\beta = -0.09$, $SE = 0.03$, $t(264.8) = -3.58$, $p < .001$, and a significant interaction between target task and prime type, $X^2(2) = 6.97$, $p = .031$. Similarly, the model with data from Experiment 5a and 5b (controlling for the target task) also showed a significant interaction between prime task and prime type, $X^2(2) = 7.94$, $p = .019$, and word-primed word evoked significantly shorter RTs than the unprimed condition, $\beta = -0.02$, $SE = 0.01$, $t(164.4) = -2.29$, $p = .023$. These results indicated that Experiment 5a, which used the pause detection task in both prime and target phase, was an ‘outlier’ with a different results pattern from the other three experiments.

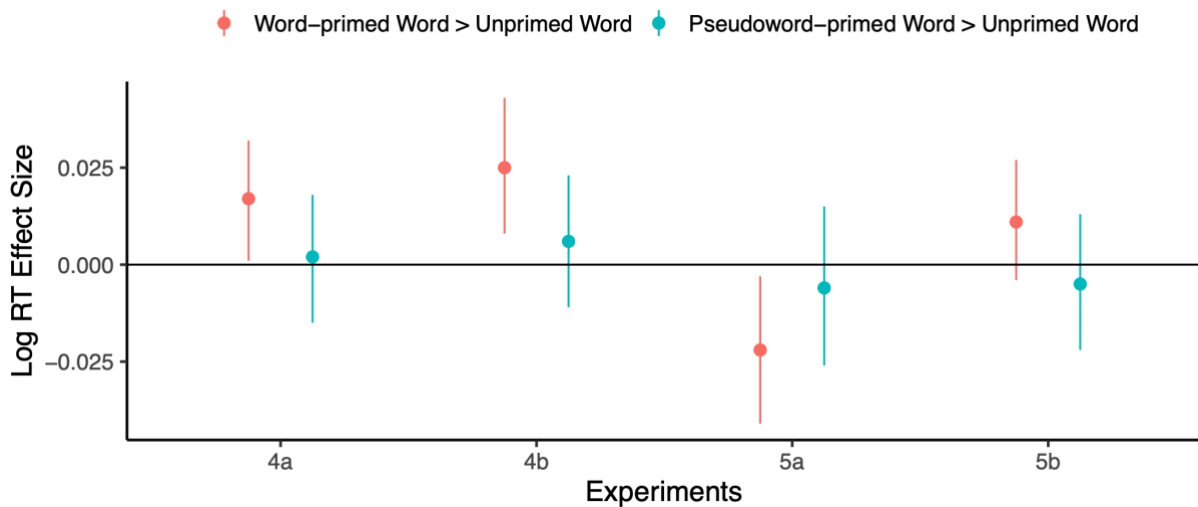


Figure 3.11. Effect size of word-primed word > unprimed word and pseudoword-primed word > unprimed word in log-transformed RT from Experiment 4a, 4b, 5a, 5b. Points show the coefficients of the effect size, error bars show the 95% credible interval, both were computed under the Bayesian framework.

3.5 General Discussion

In this Chapter, I described a series of studies that examined the role of lexical and sub-lexical processing during spoken word recognition using the competitor priming paradigm through a variety of tasks and experimental designs.

3.5.1 Competitor priming effect limited to lexical processing

As shown in Experiment 2, the competitor priming effect was robust when participants made lexical decisions during stimuli perception. This effect was not reliable when target words were primed by one or two similar sounding pseudowords. To test whether the lexical competitor priming effect can be observed when the task focuses on phonological processing, we conducted Experiment 3 using a similar paradigm as previous studies but changed the task to pause detection. The findings showed a surprising pseudoword priming effect on word targets while word competitor priming did not reliably delay word recognition as expected.

These results suggested that the competitor priming effect may not be observed during a task that focuses on sub-lexical processing even though pause detection has been proved to be sensitive to lexical activity (Mattys & Clark, 2002). However, this could also be because the pause detection task was not as effective a training task as the lexical decision task, and hence one presentation of the prime item might not be enough to induce the competitor priming effect as observed with lexical decision.

In addition, the significant delay in word recognition caused by pseudoword priming also indicated that pseudowords could be added to the lexicon and induce lexical competition not long after their initial presentation (contrary to Dumay & Gaskell, 2007; Gaskell & Dumay, 2003). Nevertheless, this result might have been false positive, since it was not replicated by Experiment 5a which also used the pause detection task in a similar design, as summarised below.

In order to better understand the findings of Experiment 3 and to dissociate the role of lexical and sub-lexical processing during spoken word recognition, we conducted Experiment 4a, 4b, 5a and 5b with separate prime and target phases, multiple presentations of each prime item and mixed the lexical decision and pause detection tasks in both phases. As mentioned above, when using the pause detection task in both prime and target phases (Experiment 5a), we did not find delayed recognition of word targets caused by pseudoword priming as shown in Experiment 3. Instead, a phonological facilitatory effect was shown for the recognition of word-primed words. On the other hand, when using the lexical decision task for both prime and target phases (Experiment 4b), the competitor priming effect on word targets remained significant although with a smaller effect size, possibly due to the long lag between prime and target items. These results seemed to show that the competitor priming effect was observable only when participants intentionally focused on lexical-level processing.

Such findings could be linked to the “good enough” comprehension process, which was found in syntactic processing literature (Ferreira & Patson, 2007), i.e. instead of always compute the complete and complex meaning of sentences, listeners fall back on an interpretation that might be inaccurate or incomplete when the task does not have such requirements. Similarly, since the pause detection task does not require participants to process the semantic meaning of items, it may be the case that lexical representations were not even thoroughly accessed by participants during the prime task, which was why no lexical competition or lexically informed prediction error was evoked during the recognition of the target word.

3.5.2 Lexical and sub-lexical processing during spoken word recognition

The competitor priming effect measures word probability (frequency) change during spoken word recognition. This effect is predicted by spoken word recognition models in different ways.

In TRACE (McClelland and Elman, 1986), frequency effect is intrinsic to lexical nodes that compete within the same level, and it has early influence on lexical access (Allopenna et al., 1998; Dahan et al., 2001). Hence, prior presentation of the competitor word enhanced its own frequency and lexical activation while inhibited the recognition of the target word. Although, according to TRACE, phoneme-level processing also affects lexical level processing through interaction, the facilitatory effect of coherent phoneme priming of the pre-DP segment may be masked by the inhibitory lexical competitor priming effect during word recognition (Vitevitch & Luce, 1998).

The Predictive Coding account (Davis & Sohoglu, 2020) proposes that lexical frequency modulates the change of connection weights between words and their constituent speech sounds (i.e. lexical to phonological connections). Hence, priming of the competitor word strengthened the prediction between pre-DP and post-DP segments via top-down connections from lexical units of the prime word. Therefore, large prediction error was evoked when such prediction was disrupted by the unexpected post-DP segments of the target word. On the other hand, when prime was presented with an inserted pause, the prediction between pre-DP and post-DP segment was temporarily disturbed, but the processing of the separated pre-DP segment was unaffected and this temporarily independent segment could still lead to predictions of multiple words beginning with the same sounds. Since the target word also shared the same pre-DP segment as the prime, its response could be facilitated by the priming of this phoneme sequence, as shown in the faster response produced by word-primed words in Experiment 5a. These results are also consistent with the reduced pre-DP MEG signals evoked by competitor primed words as shown in Chapter 2.

These opposite effects caused by lexical and sub-lexical processing are also accounted for by the Distributed Cohort model (DCM; Gaskell & Marslen-Wilson, 1997). According to the DCM, phonological and semantic representations are blended in the same nodes. Therefore, lexical competition caused by priming of the competitor word and phonological facilitation caused by priming

of the coherent pre-DP segments (when pause separated them from the post-DP segments) can be directly reflected in this distributed model structure (Gaskell & Marslen-Wilson, 2002). In Experiment 2 and 4b, when the lexical decision task was used for both prime and target phases, priming of the competitor word caused greater intrinsic competitions between lexical representations of the prime and the target word. However, in Experiment 5a, when the pause detection task was used for both phases and that each prime word was presented multiple times with and without pause, the pre-DP segment was processed temporarily as isolated ambiguous phoneme sequence which matched with multiple activated lexical items. Therefore, the coherence between the pre-DP segments of the prime and target words facilitated pause detection in the target phase.

It should also be noted that, in Experiment 4a and 5b, the lexical decision task and the pause detection task were used in prime and target phases in a mixed manner and the effects on word targets in these experiments were not reliable although there was a trend for delayed response. However, the combined results controlling for the prime or the target task showed significant interaction effect between target task and prime type, which differentiated Experiment 5a from Experiment 4a and 5b, while such difference was not found between Experiment 4b and Experiment 4a and 5b. These findings indicated that the facilitatory effect from sub-lexical processing was observable only when pause detection was involved in both prime and target phases, while the pattern of the competitor priming effect was shared by the other three experiments that all used the lexical decision task at some point during the study. These findings on word targets reflected influence from task-evoked decisions during spoken word recognition. Since, in order to make a lexical decision, participants need to use their knowledge of both word semantics and word frequency (Balota & Chumbley, 1984), while the pause detection task only requires that participants to spot whether there is an unnatural gap between phonemes. Therefore, even though the pause detection task is also sensitive to the change of lexical activity, it is the phonotactic probability that plays the most important role in making this response.

3.5.3 Facilitatory effect from pseudoword priming

As for pseudoword targets, pseudoword priming using the lexical decision task led to reliably shortened responses compared to unprimed pseudowords in both Experiment 4b and 5b.

In Experiment 4b, the lexical decision task was used for both the prime and target phases and the same lexical response was made four times for each stimulus during the prime phase. Hence, this effect was possibly due to associative learning between the initial segments of stimuli and lexical decision, which made it easier for participants to respond to target pseudowords beginning with the same initial segments. Such facilitatory task effect could also be a possible reason for the smaller effect size of the competitor priming effect on word-primed words in this experiment. However, at the same time, it was also harder for participants to respond to pseudowords that were primed by words, which were associated with the “word” response rather than the “pseudoword” response during lexical decision in the prime phase. This could explain why word-primed pseudowords were recognised more slowly than unprimed ones here.

However, in Experiment 5b, the prime and target tasks were different. Therefore, even if participants managed to form strategic association between stimuli and responses due to the lexical decision task in the prime phase, it would not be helpful for the pause detection task in the target phase. Instead, pseudoword priming with a focus on lexical processing (lexical decision task) may have activated closely resembling lexical representations that shared the same initial segments, since the pseudoword response was made by establishing that it was distinct from all known words (which was different from identifying the exact form of a single word when making the word response). Hence the phonotactic probability of the pre-DP segments was increased, which led to sub-lexical facilitatory effects on pseudoword targets (using pause detection). These results are also in line with the pre-DP results found in our MEG study, which showed reduced pre-DP neural response in pseudoword-primed items. Therefore, the faster RTs here may also indicate an advantage for sub-lexical processing of speech sounds that match multiple activated words.

3.5.4 Conclusion

This chapter described a series of experiments that studied lexical and sub-lexical processing during spoken word recognition. By manipulating the prior probability of the competitor word through the competitor priming paradigm and using tasks that focused on lexical processing (lexical decision) and sub-lexical processing (pause detection) respectively, we found that 1) the competitor priming effect was observed only when participants' focus was intentionally on lexical processing; 2) inhibitory lexical competitor priming effect and facilitatory phonological priming effect can be dissociated using the same competitor priming paradigm and a combination of different tasks. Such opposite effects of lexical and sub-lexical processing are also supported by most spoken word recognition models. Even though some of the results reported in this chapter are not completely clear-cut, they provide valuable insights on functional mechanisms of spoken word recognition.

The effect of pseudoword priming on pseudoword targets also makes me reflect more on word learning. I will discuss this further in the next chapter, which focuses on memory for spoken words and pseudowords.

4 SPOKEN WORD ENCODING – ANALYSES ON MEMORY DATA FROM EXPERIMENTS 1 - 5

4.1 Introduction

On top of effortless speech comprehension in daily life, human adults can also learn and adopt new words, such as *vlog*, *Brexit*, into their vocabulary easily (though perhaps not comfortably). Recently, one mechanism that has been recognised as important to memory encoding is the computations of the difference between predictions and perceived information, i.e. the prediction error (see Reichardt et al., 2020 and Quent et al., 2021 for review). According to the probabilistic PIMMS framework (The Predictive Interactive Multiple Memory Systems, Henson & Gagnepain, 2010), prediction error drives synaptic change (learning) between higher-level predictions and lower-level sensory signals, hence items that evoke greater prediction error should be better remembered. As reviewed in Chapter 1, previous studies using word repetition priming (Tulving & Kroll, 1995) and cloze probability (Corley et al, 2007; Haeuser & Kray, 2021) have demonstrated that large prediction error improves word memory. However, other studies (Alba & Hasher, 1983; Cycowicz et al., 2008; Höltje et al., 2019) also showed that information that is semantically congruent with schema (i.e. low prediction error) can be better encoded in memory, while irrelevant or incongruent items are remembered less well (Sweegers et al., 2015).

In this chapter, we tested whether computations of prediction error enhance memory encoding using additional recognition memory data collected as part of the experiments reported in Chapter 2 and 3. We measured the accuracy with which participants can distinguish previously heard target items from foils in a surprise subsequent memory task following the

earlier encoding phase (lexical decision/pause detection). By using the competitor priming paradigm in the encoding phase, in which the target word (e.g. *hygiene*) was primed by a similar sounding competitor word (e.g. *hijack*) sharing the same initial segments, we examined whether computations of prediction error that caused delay in spoken word recognition can also drive word memory encoding.

Mechanisms for word memory encoding during competitor priming

Based on the PIMMS account, competitor-primed words should be better remembered than unprimed ones because they evoke greater prediction error. According to PIMMS, hearing the prime word (e.g. *hijack*) enhances the prior probability (under the Bayes rule) of the association between the initial stem of the word (e.g. /haidʒ/) and the ending segment (e.g. /æk/), hence when the target item with the same initial segment but a different ending (e.g. /i:n/) is presented, prediction error should be evoked due to the disruption of the prior prediction based on the initial stem (see Figure 4.1 for an illustration of the PIMMS account applied to competitor priming). This prediction error should enhance the synaptic exchange between the medial temporal lobe (MTL) supporting the episodic memory system and the perirhinal cortex supporting the semantic memory system, which improves encoding and recollection of the target item. However, when the prime or target is a pseudoword, the prediction error should be much smaller than when both the prime and target items are words. This is because pseudowords are novel and lack prior lexical-semantic and phonological representations, hence pseudoword-based predictions or pseudoword perception are high in uncertainty in the semantic memory system, which prevents effective calculations of prediction error that drives memory encoding. Note that this is different from maximal prediction error evoked by pseudowords in the perceptual systems in the context of spoken word recognition (see Chapter

2), since pre-existing representations are especially important in forming the interactions between episodic and semantic memory systems, which is crucial for memory encoding and recollection. Overall, PIMMS predicts that one will encode word-primed words better than unprimed words, but the encoding is not as strong when a pseudoword is primed by a word and vice versa, and learning is negligible when both the prime and the target are pseudowords.

In addition, by changing the number of presentations of the prime item (four times in Experiment 4-5 vs one time in Experiment 1-3) and using tasks that focused on lexical processing (lexical decision) and sub-lexical processing (pause detection) respectively, we manipulated the strength of prediction (Greve et al., 2017) and levels of processing (Craik & Lockhart, 1972; Craik & Tulving, 1975) during memory encoding. Presenting the same item multiple times increased the certainty and the strength of prediction between the initial sounds and the ending segments of the prime item, hence the corresponding prediction error evoked by the target item should also be larger and better memory performance is expected. In addition, according to the levels of processing account, words processed at the semantic level (lexical decision) should be better encoded than words processed at the sub-lexical level (pause detection), since semantic processing is a “deeper” level of processing that produces stronger memory trace. This hypothesis is also supported by neural imaging studies using fMRI (Wagner et al., 1998b) and PET (Kapur et al., 1994), which showed stronger neural activation in the left prefrontal region for semantic processing as opposed to non-semantic processing.

Another reason why semantic processing leads to better subsequent memory is that it requires more attentional resources (i.e. more challenging or attention demanding) than the shallower sub-lexical processing (see Craik et al., 1996; Fletcher et al., 1995 for the relationship between attention and memory performance). This is also reflected by the longer response times evoked during the lexical decision task compared to the pause detection task (see Chapter 3 results). However, Jurica & Shimamura (1999) further proposed that more attentional

resources spent on processing the item may lead to worse episodic encoding of item and context association, namely there is an “item-context trade-off”. Hence shorter response times (less attentional resources) spent on the prime task may instead lead to better associative memory between the item and the context in which the item is encoded (i.e. the experimental context in which participants responded to each stimuli in this case).

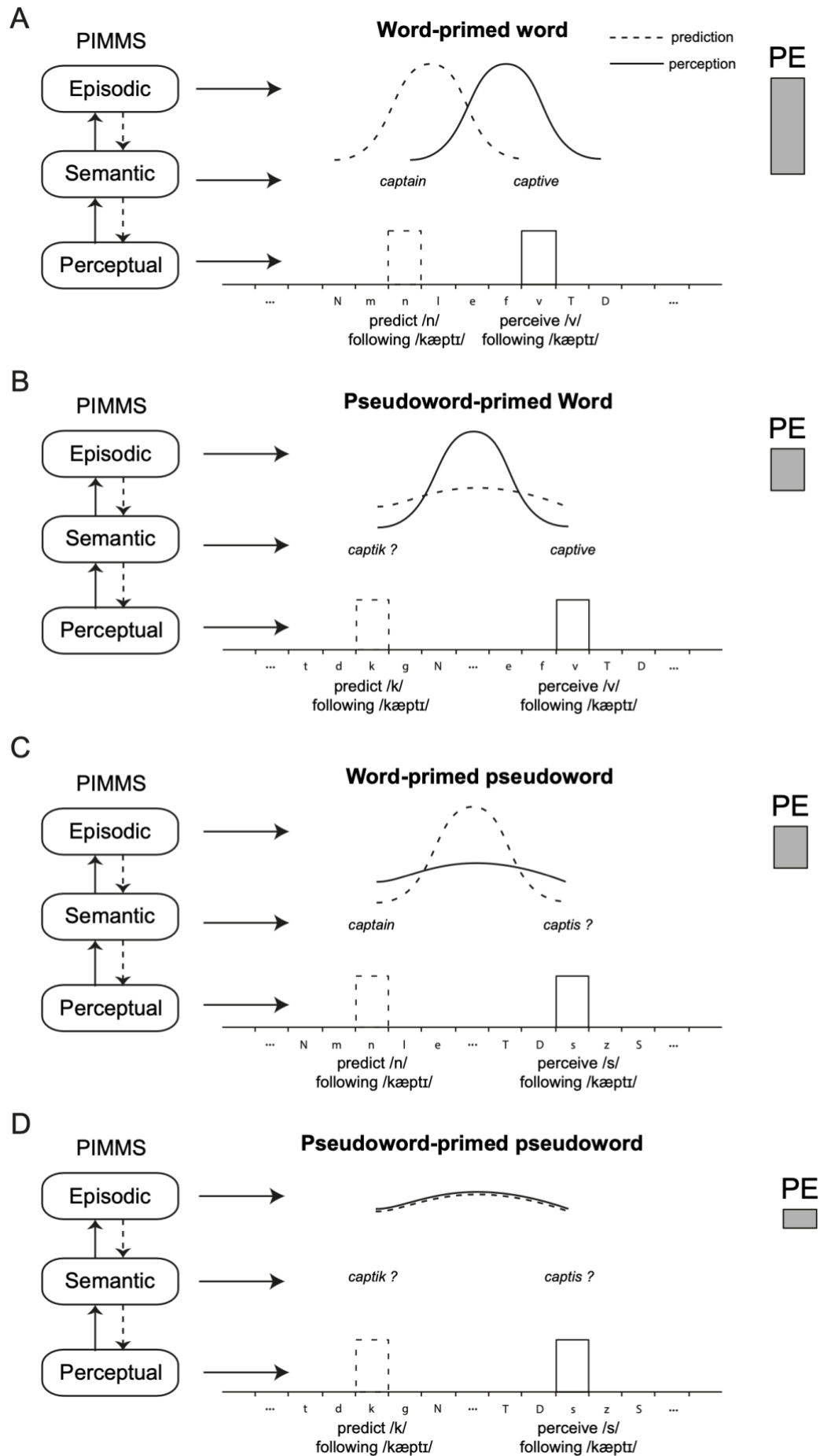


Figure 4.1. Schematic illustration of the PIMMS account with the perceptual, semantic and episodic memory systems applied to the competitor priming paradigm. Example items *captain*, *captive*, *captik* and *captis* are used as the prime word, target word, prime pseudoword, target pseudoword respectively. They share the same initial segment /kæptɪ/ (pre-DP segment), while their post-DP segments are respectively /n/, /v/, /k/, /s/. The x-axis at the perceptual system level captures similarity between phoneme features as perceived by listeners represented in phonDISC transcription. The bars in dashed lines represent the predicted post-DP segments for the target word after hearing the prime word (for simplicity, only the primed segment is shown as the predicted segment here), while the bars in solid lines shows the perception of the actual post-DP signal input of the target word. The items at the semantic memory system level shows the matching existing representations of the predicted and perceived items. Words have matching phonological, lexical and semantic representations while pseudowords do not. The distribution curves shown at the episodic memory system level reflect the certainty of prediction- and perception-based activity built on interactions between the episodic and semantic memory systems. Pointy curves reflect higher certainty than flat curves. The PE bar at the right side of each condition shows the divergence between the prediction and perception.

A. the word-primed word condition. Based on the prime word, *captain*, there is a strong prediction for /n/ after hearing /kæptɪ/. However, the sensory input in the target word after /kæptɪ/ is /v/, namely the word *captive*. Hence a large prediction error is evoked. **B.** the pseudoword-primed word condition. After hearing the prime pseudoword *captik*, the prediction for /k/ is not particularly strong due to a lack of prior phonological and lexical-semantic representations in the semantic memory system. Therefore, there is high uncertainty in the prediction. The perception of the target word *captive* should still be high in certainty as in panel A. Hence the prediction error is smaller than when both the prime and the target are words. **C.** the word-primed pseudoword condition. Like in panel A, the prediction for /n/ after hearing

/kæptɪ/ is strong. However, when perceiving the /s/ sound following /kæptɪ/ in the target item, participants encounter a novel phonological combination that does not have pre-existing representations. Hence there is high uncertainty in this perception. The prediction error is therefore also smaller than when both the prime and the target are words. **D.** the pseudoword-primed pseudoword condition. As explained for panel B and C, pseudoword-based prediction and perception are high in uncertainty due to the lack of prior phonological and lexical-semantic representations, therefore this condition evokes the smallest prediction error for memory encoding.

Neural underpinnings of word memory encoding

Apart from testing the mechanisms of memory encoding using the behavioural data, in order to further examine the neural representations of remembered and forgotten words and pseudowords in the brain, we also analysed the MEG data based on the behavioural subsequent memory performance in Experiment 1 (see Paller & Wagner, 2002, for review on subsequent memory analyses of neural data).

The PIMMS account does not predict strong pseudoword encoding due to the lack of previous knowledge about pseudowords in semantic memory system. However, it is also true that we learn novel words quite often in our daily life. One possible explanation of how this process of novel word learning happens is provided by the complementary learning systems account (CLS, McClelland et al., 1995; see Davis & Gaskell, 2009 for review). This account proposes two complementary systems for lexical acquisition, in which rapid initial acquisition of novel words (pseudowords) is supported by medial temporal and hippocampal learning (Breitenstein et al., 2005; Davis et al., 2009; Mestres-Misse et al., 2008; Takashima et al., 2014), while slow overnight consolidation in the neocortex is needed to further encode and lexicalise

novel words (Davis et al., 2009; Dumay & Gaskell, 2007; Gaskell & Dumay, 2003; see James et al., 2017, for review). However, since our MEG experiment only tested memory effects immediately after encoding, our subsequent memory analyses of the MEG data could only provide evidence for the first stage of novel word learning as described in the CLS.

In addition, previous subsequent memory studies also showed mixed results of neural underpinnings for word and pseudoword encoding. For example, Clark and Wagner (2003) showed in their fMRI study that neural activities in the left inferior prefrontal cortex (LIPC) and parietal regions predicted participants' subsequent memory for both (visually encoded) words and pseudowords although the effect was greater in posterior LIPC for pseudowords than familiar words. However, Otten and colleagues (2007) showed in their electrophysiological study that the brain supported the encoding of visually presented familiar words at frontally distributed locations from around 600ms post onset with positive going EEG signals. Whereas for pseudowords, the neural activities that supported encoding were more widely spread in the brain and happened at 1000ms onwards with negative going signals. It seems that Otten et al. (2007) showed distinct patterns of EEG signals that formed memory for words and pseudowords, while Clark and Wagner (2003) showed more consistent brain activities that corresponded to word and pseudoword memory encoding. It is therefore timely to extend the previous evidence by further investigating the neural mechanisms supporting spoken word learning.

4.2 Memory Data Analyses (Experiment 1-5)

To investigate the role of prediction error on memory encoding, we tested participants subsequent memory of target words and pseudowords presented during the encoding phase of Experiment 1 to 5 (i.e. the phase in which word perception was tested using the lexical decision

task or pause detection task, as reported in Chapter 2 and 3) using a two-alternative forced choice (2AFC) task. These memory tasks were conducted immediately after the encoding phase, with words tested before pseudowords (Figure 4.2A). When testing memory for words, we presented a semantic foil close in meaning (e.g. *sink*) with each target word (e.g. *basin*), as shown in Figure 4.2B. For pseudowords, we used phonological foils that have the same initial sounds as target pseudowords (Figure 4.2C). The reason why we used different types of foils for words and pseudowords is that it was hard to find similar sounding word phonological foil for target words when prime words were already used as the stimuli. In addition, we used the 2AFC test rather than the yes/no recognition test in order to reduce biased responses, especially because participants were likely to think that they had heard the pseudoword phonological lure, which sound very similar to the targets. More details of the 2AFC task are described in the Methods section below.

The memory tasks and data analyses of Experiments 4a, 4b, 5a and 5b were preregistered on OSF (Experiment 4a, [10.17605/OSF.IO/9453V](https://osf.io/9453V/); 4b, [10.17605/OSF.IO/JCUAR](https://osf.io/JCUAR/); 5a & b, [10.17605/OSF.IO/547GK](https://osf.io/547GK/)). While it was mentioned in the preregistration that Experiment 4a and 4b would be analysed using the frequentist approach and Experiment 5a and 5b would be analysed using the Bayesian approach, we later decided that it was most appropriate to use the frequentist approach for all the analyses of the memory data. This was because the memory results of Experiments 1-3 (see section 4.2.2) were not consistent enough and were not well supported by previous literature, hence it did not make sense to use the Bayesian framework with priors to analyse memory results of Experiment 4-5.

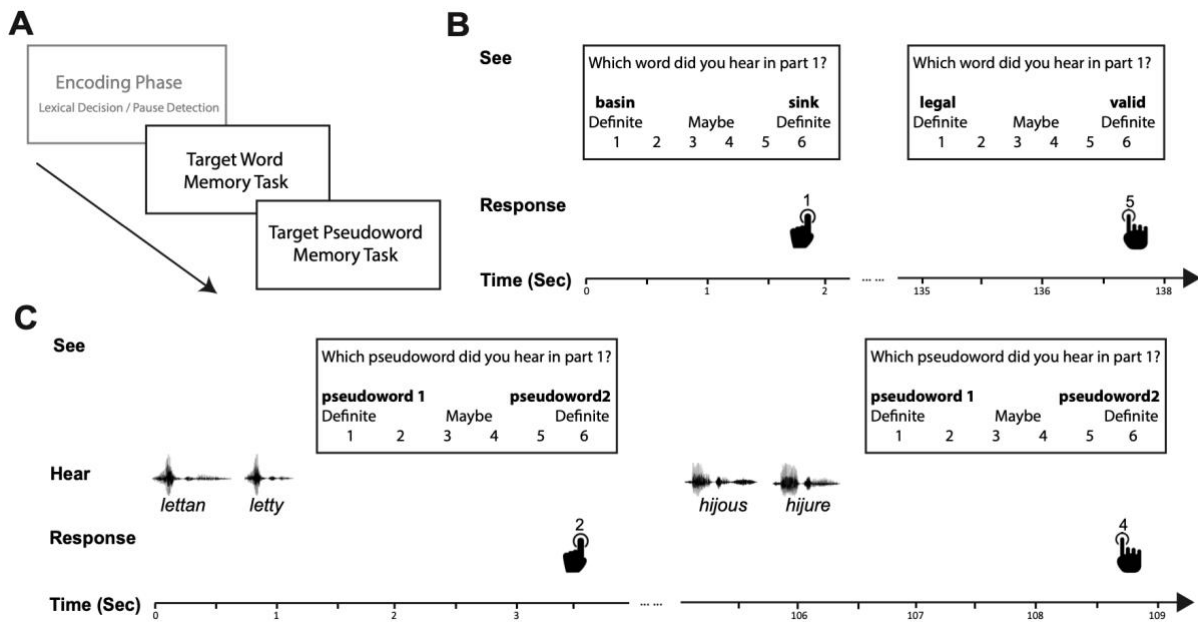


Figure 4.2. The paradigm of memory tasks used in Experiment 1 to 5. **A.** The memory tasks for target words and pseudowords came after the encoding phase as reported in Chapter 2 and 3. **B.** Word memory task. Participants saw two words (one is the target, the other is the semantic foil) on the screen each time and chose the word that they remembered hearing in the encoding phase while indicating their confidence by pressing a key. **C.** Pseudoword memory task. Participants heard two pseudowords each time (one is the target, the other is the phonological foil) and chose the one that they remembered hearing in the encoding phase while indicating their confidence by pressing a key.

4.2.1 Methods

Participants

As the memory tasks were conducted immediately after the encoding phase, in which participants performed word and pseudoword perception tasks (in the form of lexical decision or pause detection task) in each experiment, all data described in this chapter were collected

from the same participants as mentioned in Chapter 2 and 3. Participants removed from the encoding phase due to high error rates (two standard deviations above the mean) were also removed from the memory analyses. Similarly, participants with high error rates in the word memory task were also removed from our analyses. In addition, due to the greater difficulty of the pseudoword memory task, participants with high probability of random responding in this task (with lower than 80% cumulative probability of observing correct responses in up to half of all trials in a binomial test as the threshold for Exp 1-4, and 60% as the threshold for Exp5a&b due to longer lags between prime and target items in these experiments) were excluded from further analyses. As a result, we included 22 participants in Experiment 1, 31 participants in Experiment 2, 51 participants in Experiment 3, 72 participants in Experiment 4a, 70 participants in Experiment 4b, 92 participants in Experiment 5a, 84 participants in Experiment 5b.

Experimental Design and Materials

The memory tasks were conducted following the encoding phase (i.e. word and pseudoword perception in the form of lexical decision or pause detection task) in each experiment (see Figure 4.2A), hence the same repeated measures design of lexicality and prime type factors as described in Chapter 2 and 3 also applied to the memory tasks. These tasks focused on episodic memory for target items (as opposed to prime items) that were presented previously during the encoding phase (in the case of Experiment 4-5, all target items were presented in the target phase, which we refer to as the encoding phase in this Chapter). In order to avoid unnecessary distraction from word and pseudoword perception during the encoding phase, participants were not informed in advance that their memory of the target items would be tested.

As shown in Figure 4.2B and C, Two-alternative forced choice (2AFC) tasks were used to test participants' episodic memory of target words and pseudowords, in which participants

were asked to choose which one out of the two items had been presented in the encoding phase and indicated confidence ratings on a 1-6 scale on the keyboard (1-4 scale on the button box in Experiment 1 due to the limited number of buttons), with 1 indicating very confident in the item shown on the left and 6 indicating very confident in the item shown on the right, 2 and 5 indicating less confidence and 3 and 4 indicating guessing. Each word target was paired with a semantic foil (close in meaning, but not necessarily sounding similar, e.g. *sink* as a semantic foil for the target word *basin*) and was presented visually on the screen together with the foil. Each pseudoword was paired with a phonological foil (close in pronunciation and starting with the same initial segments, e.g. *hijous* as a phonological foil for the target pseudoword *hijure*) and was presented auditorily only, since participants never saw what each pseudoword looked like during the encoding phase.

In addition, since the number of participants was relatively small in the MEG study (Experiment 1), we added a second phase of perception and memory tasks using the rest of the stimuli that we have created (see Figure 4.3, e.g., if *hijack* and *hygiene* were used in the first phase as the prime and target word, then *higent* and *hijure* sharing the same initial segments were used in the second phase as the prime and target pseudoword) so as to maximize the power of the experiment, especially for the subsequent memory analyses of MEG data as described in section 4.3. Participants went through the same perception and memory tasks in the second phase of the experiment. We describe the behavioural memory results of data from phase 1, phase2 and both phases of Experiment 1 in the results section 4.2.2. Note that results from phase 2 might differ from phase 1 due to the fact that memory tests in phase 2 were no longer surprise tests and that participants might change the way they encoded items in phase 2. Another caveat is that “unprimed” items in phase 2 were not truly unprimed, which is also why word perception results from the second phase are not reported in Chapter 2.

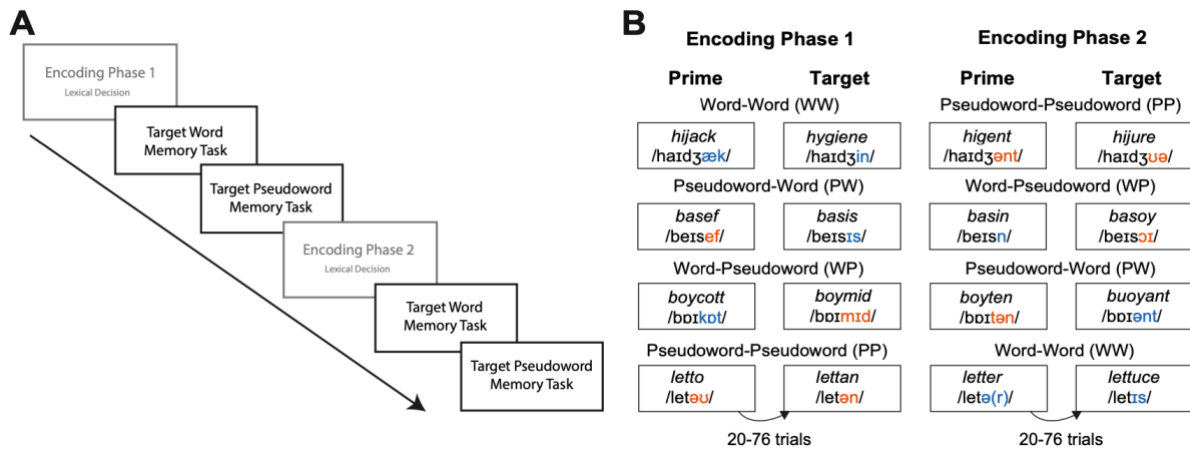


Figure 4.3. The paradigm and design of Experiment 1 (the MEG study). A. The full experiment consisted of two phases of encoding and memory tasks. B. The lexical decision task (encoding) from both phase 1 and phase 2 involved prime and target items sharing the same initial segments and the same four conditions (word-primed word, pseudoword-primed word, word-primed pseudoword and pseudoword-primed pseudoword). Each item was presented once only in the entire experiment.

Data analysis

Response accuracy and confidence ratings of the memory tasks were measured. While accuracy results simply reflected whether participants have remembered or forgotten an item, confidence rating results were more sensitive to how well an item was remembered or how completely an item was forgotten. The data were analysed using mixed effect models, with lexicality and prime type as fixed factors and participants and items as random factors. Accuracy was analysed using logistic mixed-effect model with the *lme4* package and confidence rating (ordinal data) was analysed using the mixed-effect cumulative link model with the ordinal package (Christensen, 2015) implemented in R. Similar to analyses for word and pseudoword perception data from the encoding phase, we attempted to run maximal models with full hierarchical structures wherever possible, but reduced random effects structures were applied

when the full model did not converge. Likelihood-ratio tests comparing the full model to a nested reduced model using the Chi-Square distribution were conducted to evaluate the lexicality effect and the interaction effect between lexicality and prime type. Significance of model coefficients were obtained using z statistics from the model summary. Since we were interested in contrasts between the unprimed condition and each primed condition and we did not have definite *a priori* predictions, Bonferroni correction was applied to these pairwise contrasts. We excluded target items that were not responded to correctly during the encoding phase from further analyses in the memory task. See Appendix D for descriptive statistics for the data presented in all figures in this Chapter.

4.2.2 Results & Discussion

Experiment 1

We analysed the results from phase 1 (Figure 4.4A&B), phase 2 (Figure 4.4C&D) and from both phases (Figure 4.4E&F) of Experiment 1 respectively. The descriptive statistics for the data presented in the plots are reported in Appendix D.

Phase 1. Results from phase 1 only showed significant effect of lexicality – memory accuracy and confidence rating for pseudowords were worse than words, $X^2(3) = 37.59$, $p < .001$ (accuracy), $X^2(3) = 41.79$, $p < .001$ (rating). However, there was no reliable interaction between lexicality and prime type on either accuracy, $X^2(2) = 0.66$, $p = .720$, or rating, $X^2(2) = 0.58$, $p = .750$. In addition, the main effect of prime type was not significant for either words, $X^2(2) = 3.88$, $p = .144$ (accuracy), $X^2(2) = 2.40$, $p = .302$ (rating), or pseudowords $X^2(2) = 1.88$, $p = .391$ (accuracy), $X^2(2) = 3.06$, $p = .217$ (rating).

Phase 2. Results from phase 2 also showed significant lexicality effect on accuracy, $X^2(3) = 41.54, p < .001$, and rating, $X^2(3) = 40.65, p < .001$, while there was no interaction between lexicality and prime type, $X^2(2) = 1.87, p = .392$ (accuracy), $X^2(2) = 3.55, p = .170$ (rating). However, different from phase 1, there was a main effect of prime type for both word and pseudoword target item memory accuracy, $X^2(2) = 9.47, p = .009$ (words); $X^2(2) = 12.62, p = .002$ (pseudowords), and for pseudoword rating, $X^2(2) = 11.41, p = .003$. More specifically, there was a trend that pseudoword priming disrupted word memory, $\beta = -0.39, SE = 0.19, z = -2.08, p = .076$, whereas pseudoword memory was enhanced by pseudoword priming, as reflected in both accuracy, $\beta = 0.41, SE = 0.16, z = 2.50, p = .025$, and rating, $\beta = -0.25, SE = 0.09, z = -2.91, p = .007$.

Both Phases. When using data from both phase 1 and phase 2, we again saw the significant lexicality effect for both accuracy, $X^2(3) = 42.25, p < .001$, and rating, $X^2(3) = 44.76, p < .001$, and there was no significant interaction effect between lexicality and prime type, $X^2(2) = 0.75, p = .687$ (accuracy), $X^2(2) = 3.7, p = .157$ (rating). In addition, the pairwise effects were similar to the results from phase 2 alone, with marginally worse memory accuracy for pseudoword-primed words compared to unprimed words, $\beta = -0.30, SE = 0.14, z = -2.07, p = .077$, and better memory for pseudoword-primed pseudowords compared to unprimed pseudowords in terms of both accuracy, $\beta = 0.26, SE = 0.11, z = 2.29, p = .044$, and rating, $\beta = -0.14, SE = 0.06, z = -2.54, p = .022$.

We also tested the effect of phase on target words and pseudowords separately. In word models, we found no interaction effect between word conditions and phase, $X^2(2) = 1.18, p = .556$ (accuracy), $X^2(2) = 1.28, p = .527$ (rating), or main effect of phase, $X^2(3) = 1.65, p = .649$ (accuracy), $X^2(3) = 2.10, p < .553$ (rating). Whereas in pseudoword models, the prime effect was affected by phase as shown in the marginally significant interaction by accuracy, $X^2(2) = 5.20, p = .074$, and reliable interaction by rating, $X^2(2) = 6.05, p = .048$, and the overall

pseudoword memory accuracy and ratings are also significantly worse in phase 2 compared to phase 1, $X^2(3) = 27.74, p < .001$ (accuracy), $X^2(3) = 25.96, p < .001$ (rating).

However, it is worth noting that the priming effect is less clear-cut in phase 2, because all items presented in phase 2 followed two other items sharing the same initial segments from phase 1 (see Figure 4.2B), hence participants might still remember those two similar sounding items from phase 1 when performing the tasks in phase 2. Moreover, participants were also aware that there would be memory tasks following the encoding phase in phase 2, which could potentially change their encoding or response strategy. In addition, ‘unprimed’ and ‘word-primed’ pseudowords from phase 2 also had the same foil as the pseudowords sharing the same initial segments presented in phase 1, hence having heard the same foil before could also have biased participants in making incorrect responses in these two conditions in phase 2, which could be a possible reason for the better memory in pseudoword-primed condition only.

In short, the results from phase 2, especially those of pseudoword targets, are potentially confounded by multiple factors on top of the intended priming design, while results from phase 1 alone could be under powered. Therefore, we included the memory tasks in our subsequent experiments as well to further examine the effect of competitor priming on word and pseudoword memory. Additionally, the MEG data from both phases of this study were also analysed for subsequent memory effects and are reported later in section 4.3 of this chapter.

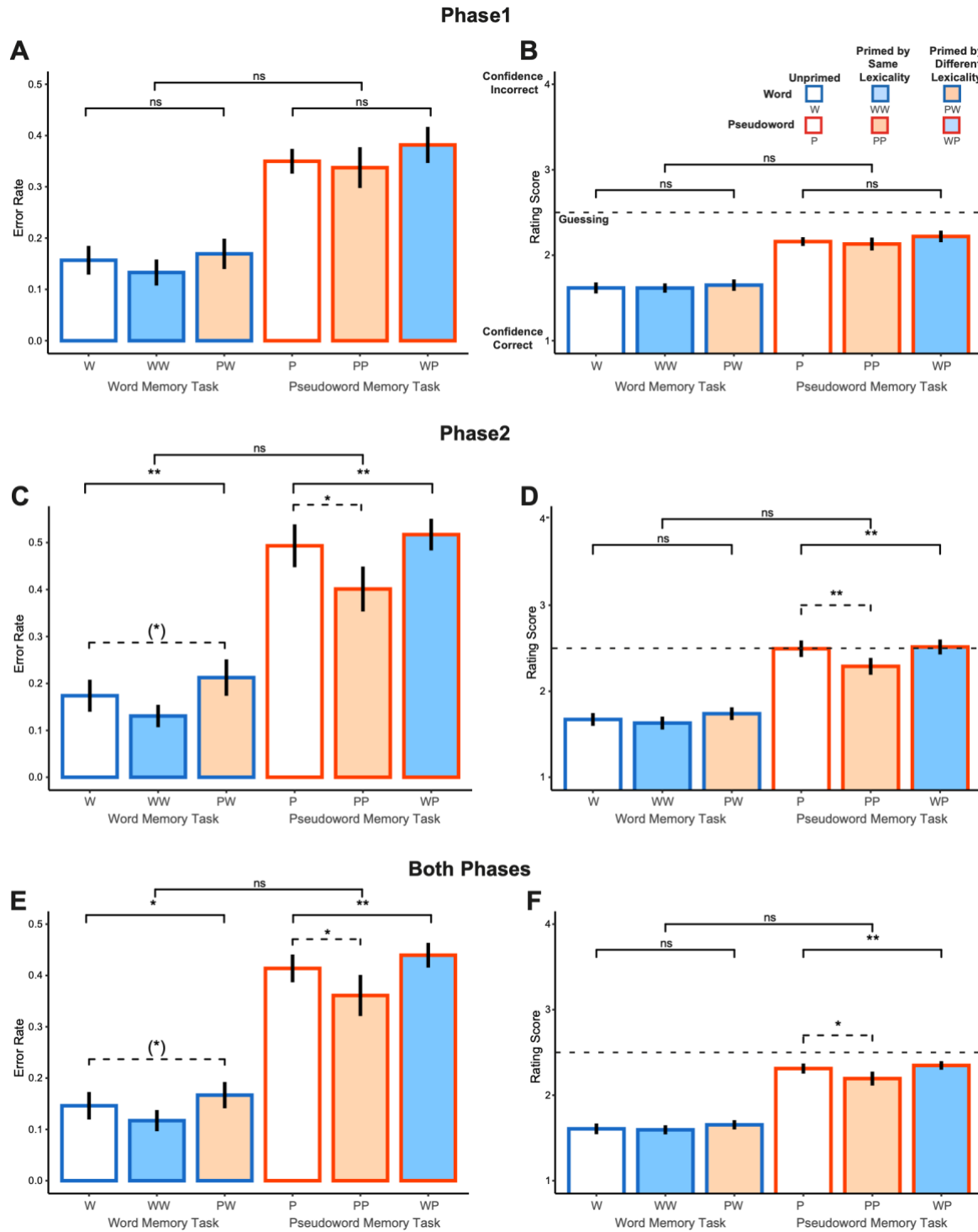


Figure 4.4. Accuracy (panel *A, C, E*) and confidence rating results (panel *B, D, F*) from the memory tasks from phase 1, phase 2 and both phases of Experiment 1 (the MEG study). Bars are color-coded by lexicality and prime type on the *x* axis (words, blue frame; pseudowords, orange frame; unprimed, no fill; primed by same lexicality, consistent fill and frame colors; primed by different lexicality, consistent fill and frame colors).

primed by different lexicality, inconsistent fill and frame colors). Bars show the subject grand averages, error bars represent \pm within-subject CI, adjusted to remove between-subjects variance (Cousineau, 2005). Statistical significance is shown based on generalised linear mixed-effects regression (for accuracy analyses) and ordinal mixed-effect regression (for rating analyses) in log-odds: (*) $p < 0.1$, * $p < 0.05$, ** $p < 0.01$. Statistical comparisons shown with solid lines indicate the lexicality by prime-type interaction and main effects of prime-type for each lexicality, while broken lines indicate the significance of pairwise comparisons.

Experiment 2

In this study, there was an extra condition in the encoding phase, double primed by different lexicality, (see Figure 3.2 from Chapter 3 for the paradigm) on top of the existing conditions in Experiment 1.

Figure 4.5 shows consistent results of memory accuracy and confidence rating from Experiment 2. Like in Experiment 1, memory error rate and rating were significantly higher for pseudowords than words, $X^2(4) = 77.1, p < .001$ (accuracy), $X^2(4) = 98.12, p < .001$ (rating). Moreover, there was a significant interaction between prime type and lexicality, $X^2(3) = 14.10, p = .003$ (accuracy), $X^2(3) = 25.80, p < .001$ (rating), meaning that the competitor priming effect on memory was different for words and pseudowords. Specifically, pseudoword priming disrupted word memory comparing to unprimed words, $\beta = -0.56, SE = 0.20, z = -2.82, p = .015$ (accuracy), $\beta = 0.17, SE = 0.06, z = 2.90, p = .011$ (rating). Although there was a similar trend on words that were double primed by pseudowords, the effect did not survive correction for multiple comparison. In addition, pseudoword priming also retarded pseudoword memory

comparing to unprimed pseudowords, $\beta = -0.24$, $SE = 0.10$, $z = -2.41$, $p = .048$ (accuracy), $\beta = 0.10$, $SE = 0.05$, $z = 2.19$, $p = .084$ (rating).

The results of word targets seemed consistent with those in Experiment 1, however it was unclear why double primed words did not produce a similarly, if not more, measurable effect compared to the single primed condition. On the other hand, the results of pseudoword targets were opposite to those found in Experiment 1, further studies are therefore necessary to test whether these effects are replicable.

It should be noted that the PIMMS account does not predict these decremental memory effects caused by pseudoword priming. According to PIMMS, memory performance is proportional to the size of prediction error, one prior presentation of a pseudoword should not have produced any strong prediction since pseudowords are novel items with low prior probability, hence there should not be any reliable enhancement of memory performance, but a negative memory effect was not predicted either. Additionally, the attentional resources account predicts worse associative memory between item and experimental context caused by greater attentional effort on responding to the item. However, pseudoword-primed items did not evoke any reliably longer response times than unprimed items during the lexical decision task in the encoding phase, hence their disrupted memory was not expected.

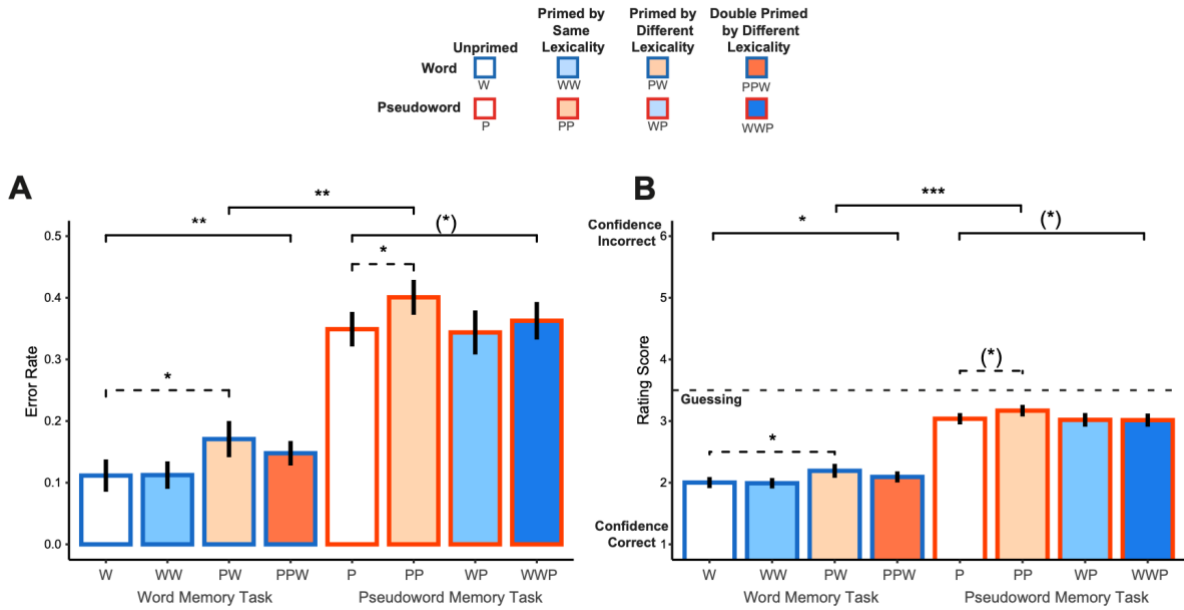


Figure 4.5. Accuracy (**A**) and confidence rating (**B**) memory results from Experiment 2. Bars are color-coded by lexicality and prime type on the *x* axis (words, blue frame; pseudowords, orange frame; unprimed, no fill; primed by same lexicality, consistent fill and frame colors; (double) primed by different lexicality, inconsistent fill and frame colors). Bars show the subject grand averages, error bars represent \pm within-subject CI, adjusted to remove between-subjects variance. Statistical significance is shown based on generalised linear/ordinal mixed-effects regression in log-odds: (*) $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Statistical comparisons shown with solid lines indicate the lexicality by prime-type interaction and main effects of prime-type for each lexicality, while broken lines indicate the significance of pairwise comparisons.

Experiment 3

This study used the pause detection task rather than lexical decision task in the encoding phase and removed the double primed by different lexicality condition.

As before, results from Experiment 3 (Figure 4.6) also indicated worse memory for pseudowords than words, $X^2(3) = 57.08, p < .001$ (accuracy), $X^2(3) = 66.83, p < .001$ (rating), and there was also a reliable interaction between prime type and lexicality, $X^2(2) = 5.52, p = .063$ (accuracy), $X^2(3) = 6.98, p = .031$ (rating). In addition, while the main effect of prime type was not significant for pseudowords, $X^2(2) = 1.10, p = .576$ (accuracy), $X^2(2) = 2.00, p = .368$ (rating), it was reliable for words, $X^2(2) = 7.69, p = .021$ (accuracy), $X^2(2) = 7.75, p = .021$ (rating), specifically both word and pseudoword priming disrupted word memory compared to the unprimed condition, $\beta = -0.23, SE = 0.09, z = -2.53, p = .023$ (word-primed word accuracy), $\beta = 0.08, SE = 0.04, z = 2.01, p = .090$ (word-primed word rating); $\beta = -0.24, SE = 0.10, z = -2.51, p = .024$ (pseudoword-primed word accuracy), $\beta = 0.13, SE = 0.05, z = 2.88, p = .008$ (pseudoword-primed word rating).

Consistent with results from Experiment 1 and 2, pseudoword priming seemed to have a negative effect on word memory regardless of the encoding task. Again, these results are not predicted by the PIMMS account. However, combined with the behavioural results from the encoding phase (see Figure 3.5 in Chapter 3), in which pseudoword-primed words evoked longer response time than unprimed words, it seems possible that the disrupted episodic memory of pseudoword-primed words could be caused by “item-context trade off”, as suggested by the attentional resources account. However, as discussed in Chapter 3, the delayed recognition of pseudoword-primed words shown in Experiment 3 was unexpected and not replicated in our subsequent studies, which makes the memory results harder to explain in terms of attentional resources. Another possible explanation is that there could be some sort of proactive interference from the prime item during memory retrieval process, as suggested by the contextual binding account of episodic memory (Yonelinas et al., 2019).

Overall, since the encoding process in Experiment 1-3 were all based on a single prior presentation of a similar sounding item, it is unclear if this disruptive priming effect is

generalizable when stronger encoding is involved, e.g. multiple presentations of the same prime.

Therefore, results from memory tests in Experiment 4-5 would help answer this question.

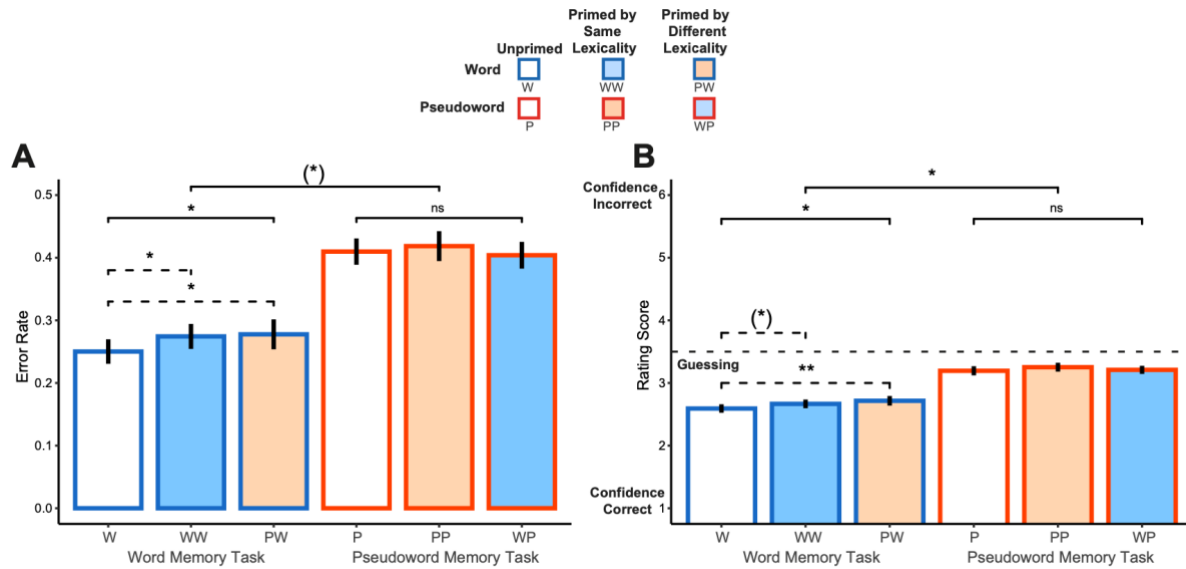


Figure 4.6. Accuracy (**A**) and confidence rating (**B**) memory results from Experiment 3. Bars are color-coded by lexicality and prime type on the x axis (words, blue frame; pseudowords, orange frame; unprimed, no fill; primed by same lexicality, consistent fill and frame colors; primed by different lexicality, inconsistent fill and frame colors). Bars show the subject grand averages, error bars represent \pm within-subject CI, adjusted to remove between-subjects variance. Statistical significance is shown based on generalised linear/ordinal mixed-effects regression in log-odds: (*) $p < 0.1$, * $p < 0.05$, ** $p < 0.01$. Statistical comparisons shown with solid lines indicate the lexicality by prime-type interaction and main effects of prime-type for each lexicality, while broken lines indicate the significance of pairwise comparisons.

Experiment 4-5

As illustrated in Figure 3.6 in Chapter 3, Experiment 4a, 4b, 5a, 5b adopted a design with separate prime and target phases instead of interleaved prime and target trials as in Experiment 1 to 3. Each prime item was presented 4 times in the prime phase, while each target item was presented once in the target phase and their memory was tested subsequently, hence the target phase is referred to as the encoding phase here. As shown in Table 3.2 in Chapter 3, Experiment 4a and 5a used the pause detection task as the prime task, while Experiment 4b and 5b used the lexical decision task as the prime task; Experiment 4a and 4b used the lexical decision task as the encoding (target) task, while Experiment 5a and 5b used the pause detection task as the encoding task. Hence, Experiment 4b and 5a were built on Experiment 1, 2 and 3 with multiple presentations of the prime, while Experiment 4a and 5b were entirely new.

Experiment 4a. Figure 4.7A&B show that memory for pseudowords were worse than that for words, $X^2(3) = 141.76, p < .001$ (accuracy), $X^2(3) = 128.13, p < .001$ (rating), and that there was no interaction effect between prime type and lexicality, $X^2(2) = 2.71, p = .258$ (accuracy), $X^2(2) = 0.24, p = .887$ (rating). In addition, accuracy results showed that memory for pseudoword-primed words was significantly better than unprimed words, $\beta = 0.45, SE = 0.20, z = 2.25, p = .048$, although this effect was not reliable in the rating results, $\beta = -0.03, SE = 0.05, z = -0.72, p = .470$. No significant effect of prime type was found for pseudoword targets.

Experiment 4b. The results of this experiment (Figure 4.7C&D) also indicated a significant difference between memory for words and pseudowords, $X^2(3) = 170.13, p < .001$ (accuracy), $X^2(3) = 151.41, p < .001$ (rating). In addition, rating results showed that the priming effect was affected by target items' lexicality, $X^2(2) = 9.09, p = .011$, while this interaction was not reliable in accuracy results, $X^2(2) = 3.38, p = .184$. Memory for words was reliably

enhanced when they were primed by another similar sounding word, this was especially robust in rating, $\beta = 0.46$, $SE = 0.23$, $z = 1.99$, $p = .094$ (accuracy), $\beta = -0.18$, $SE = 0.05$, $z = -3.64$, $p < .001$ (rating). Similar effects were also reliable for pseudoword-primed words as shown in their memory rating, $\beta = -0.14$, $SE = 0.05$, $z = -2.95$, $p = .006$. Although there was also a significant main effect of prime type for pseudowords shown in rating, $X^2(2) = 17.21$, $p < .001$, the pairwise comparison effects for unprimed pseudowords and pseudoword-primed or word-primed pseudowords did not survive correction for multiple comparison, $\beta = -0.01$, $SE = 0.04$, $z = -0.23$, $p = .817$ (pseudoword-primed vs unprimed pseudowords), $\beta = -0.07$, $SE = 0.04$, $z = -1.72$, $p = .170$ (word-primed vs unprimed pseudowords).

Experiment 5a. Apart from significant lexicality effect, $X^2(3) = 116.61$, $p < .001$ (accuracy), $X^2(3) = 137.82$, $p < .001$ (rating), no other reliable effects were found in results of this experiment (Figure 4.7E&F).

Experiment 5b. As shown in Figure 4.7G&H, the results of Experiment 5b shared a similar pattern as those of Experiment 4b. On top of the significant lexicality effect, $X^2(3) = 121.79$, $p < .001$ (accuracy), $X^2(3) = 127.26$, $p < .001$ (rating), there was also a significant interaction between lexicality and prime type, $X^2(2) = 6.40$, $p = .041$ (accuracy), $X^2(2) = 12.32$, $p = .002$ (rating). Moreover, main effect of prime type was only found for words, $X^2(2) = 11.77$, $p = .003$ (accuracy), $X^2(2) = 15.66$, $p < .001$ (rating), but not for pseudowords, $X^2(2) = 0.002$, $p = .999$ (accuracy), $X^2(2) = 0.07$, $p = .964$ (rating). Specifically, participants had significantly better memory for both word-primed and pseudoword-primed words than unprimed words, $\beta = 0.22$, $SE = 0.08$, $z = 2.70$, $p = .014$ (word-primed vs unprimed words accuracy), $\beta = -0.12$, $SE = 0.04$, $z = -3.24$, $p = .002$ (word-primed vs unprimed words rating); $\beta = 0.30$, $SE = 0.09$, $z = 3.36$, $p = .002$ (pseudoword-primed vs unprimed words accuracy), $\beta = -0.14$, $SE = 0.04$, $z = -3.81$, $p < .001$ (pseudoword-primed vs unprimed words rating).

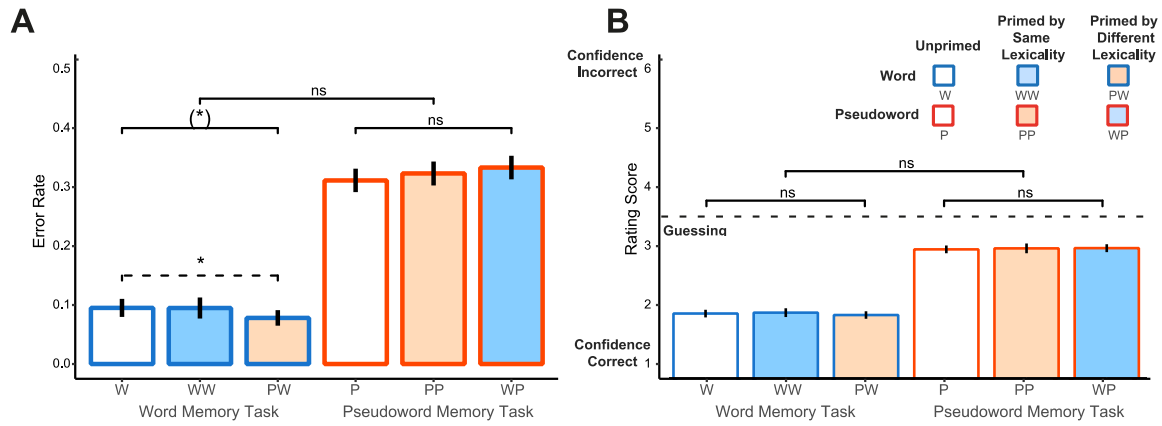
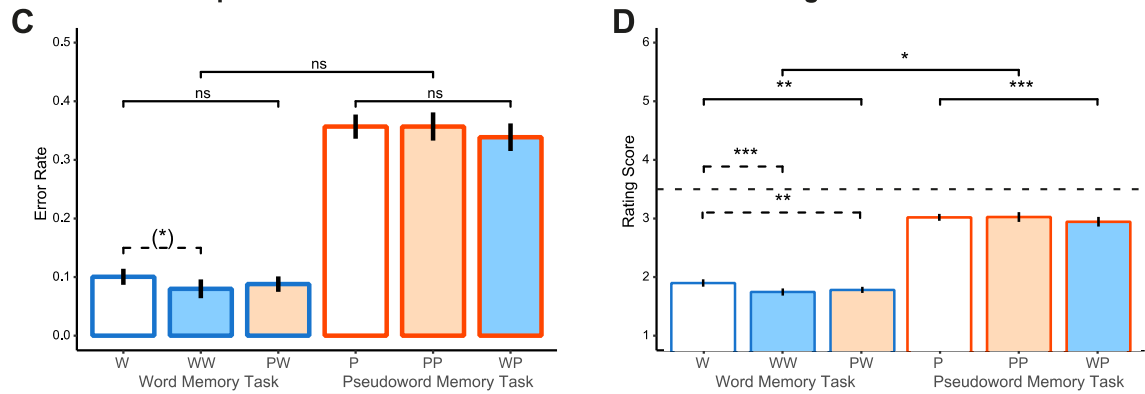
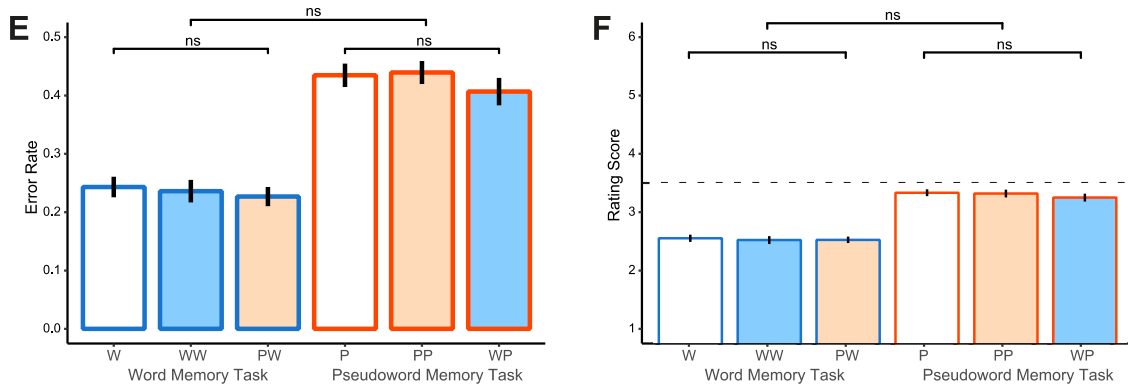
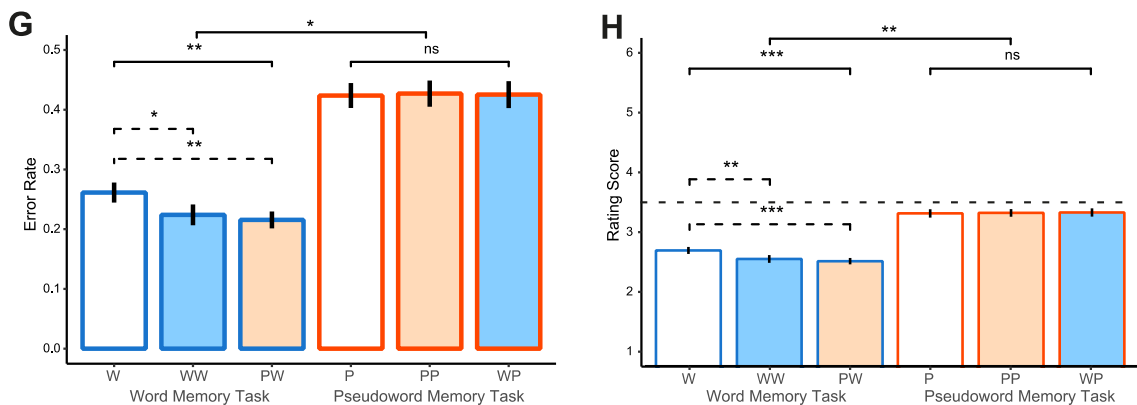
Experiment 4a Prime: Pause Detection - Encoding: Lexical Decision**Experiment 4b Prime: Lexical Decision - Encoding: Lexical Decision****Experiment 5a Prime: Pause Detection - Encoding: Pause Detection****Experiment 5b Prime: Lexical Decision - Encoding: Pause Detection**

Figure 4.7 Accuracy (*A, C, E, G*) and confidence rating (*B, D, F, H*) memory results from Experiment 4-5. Bars are color-coded by lexicality and prime type on the x axis (words, blue frame; pseudowords, orange frame; unprimed, no fill; primed by same lexicality, consistent fill and frame colors; primed by different lexicality, inconsistent fill and frame colors). Bars show the subject grand averages, error bars represent \pm within-subject CI, adjusted to remove between-subjects variance. Statistical significance is shown based on generalised linear/ordinal mixed-effects regression in log-odds: (*) $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Statistical comparisons shown with solid lines indicate the lexicality by prime-type interaction and main effects of prime-type for each lexicality, while broken lines indicate the significance of pairwise comparisons.

Comparison of word results from Experiment 4-5. In order to compare the results from these four experiments, we conducted factorial analyses on combined word data with prime type, prime task and encoding task as factors. Both the accuracy model and rating model with data from all four experiments showed a significant main effect of encoding task, namely using pause detection as the encoding task (Experiment 5a & b) led to higher error rates and lower confidence ratings compared to using lexical decision as the encoding task (Experiment 4a & b), $\beta = -1.51$, $SE = 0.16$, $z = -9.68$, $p < .001$ (accuracy), $\beta = 0.74$, $SE = 0.07$, $z = 10.12$, $p < .001$ (rating), indicating that memory encoding during the lexical decision task was more effective than during the pause detection task overall. In addition, the confidence rating model also showed significant interaction effects between prime task and word priming, $\beta = 0.20$, $SE = 0.06$, $z = 3.59$, $p < .001$, and prime task and pseudoword priming, $\beta = 0.13$, $SE = 0.06$, $z =$

2.37, $p = .018$. These results suggested that priming effects on target word memory were larger when the prime task was lexical decision than when it was pause detection.

Overall, the results of Experiment 4-5 showed that there was better memory encoding when participants performed the lexical decision task than when they performed the pause detection task in the target phase. This is consistent with the Levels of Processing model (Craik & Lockhart, 1972), with semantic processing being in a deeper level than phonetic processing, which hence contributes to stronger memory trace. In addition, the results of multiple prior presentations of the same competitor word or pseudoword leading to better target word memory than unprimed words contrasted with those of Experiment 1-3, which showed interference from single prior presentation.

However, better memory for primed words could be due to memory for the common stems shared by the prime and the target items, which were different from those of the semantic foil used in the word memory task. These effects were absent in pseudowords, as the pseudoword phonological foil shared the same initial segments as the prime and target items, hence the pseudoword memory task probed the offset of the item only. That being said, the fact that these encoding effects on words were only significant when the prime task was lexical decision but not pause detection indicated that stem encoding, if any, was not sufficient to cause a reliable difference between the memory for primed and unprimed words.

These results were not explainable by the attentional resources account that proposes “item-context trade-off”, since better remembered words in Experiment 4b and 5b did not occupy less attentional resources (shorter response times) during the encoding task (see Figure 3.8 in Chapter 3). Importantly, memory for both words and pseudowords were worse when the encoding task was pause detection compared to when the task was lexical decision, yet participants’ response times were also reliably shorter during the pause detection task, hence

less effort in performing the encoding task did not seem to have contributed enough to episodic memory that associated the item with experimental context such that it could facilitate participants' subsequent recognition memory of the items.

On the contrary, the results were in line with the PIMMS account, in which stronger prediction between syllables were generated during the prime phase that focused more on lexical semantic processing, hence the disruption of the prediction shown in target words evoked greater prediction error, which in turn enhanced encoding. It should also be noted that, apart from word-primed words, pseudoword-primed words were also reliably better remembered than unprimed words when the prime task was lexical decision. One possible explanation is that, when the same pseudoword prime was encountered multiple times rather than one time only during the lexical decision task, it also built up activation for neighbouring words which could have possibly strengthened its lexical representation. In contrast, when the prime task focused mainly on phonological processing (i.e. the pause detection task), no strong predictions between the initial and ending segments were generated for prime items, hence prediction error was not strongly evoked on target items either. These word memory results also matched with the opposite effects of lexical processing and sub-lexical processing on target word perception (see Chapter 3), in which the competitor priming effect was observable only when the lexical decision task was used. Such effects are supported by the Predictive Coding account of spoken word recognition, which also promotes the role of computations of prediction error.

However, the fact that memory effects were not shown for pseudoword targets indicated that very little prediction error was evoked by pseudoword targets due to the high uncertainty of their representations in the semantic memory system. It could also be that the phonological memory task used for pseudoword targets was too difficult, as there was no facilitation from stem encoding.

4.3 Subsequent Memory Analyses of MEG Data

In order to further understand the relationship between neural responses during spoken word and pseudoword perception and subsequent memory, we analysed the MEG data from both phases of Experiment 1 based on subsequent behavioural memory results. The conditions for analyses were divided into remembered word, forgotten word, remembered pseudoword and forgotten pseudoword. Prime types were not considered here due to the confounding effects in Phase 2 of the study and the small number of trials in forgotten conditions (see Table 4.1 below).

Through the subsequent memory analyses, we aimed to examine whether neural activity as recorded by the MEG could predict subsequent memory of the items (Paller & Wagner, 2002), the brain locations that support word and pseudoword memory encoding (Breitenstein et al., 2005; Clark & Wagner, 2003; Davis & Gaskell, 2009; Davis et al., 2009; Otten et al., 2007; Paller & Wagner, 2002) and whether remembered and forgotten words and pseudowords were represented differently in the brain (Clark & Wagner, 2003; Otten et al., 2007).

4.3.1 Methods

Participants

The same twenty-four participants as described in Chapter 2 also took part in the memory task immediately following the lexical decision task. Two participants that were removed from word recognition analyses were also excluded from the subsequent memory analyses, one further participant was excluded from the subsequent memory analyses due to technical problems, hence resulting in 21 participants in total.

Preprocessing & Data analyses

Based on response accuracy in word and pseudoword memory tasks from both phases of Experiment 1 (see Figure 4.4 from Section 4.2), we averaged neural responses on each trial during the lexical decision task based on subsequent memory performance and item lexicality, i.e. each trial was marked as remembered word, forgotten word, remembered pseudoword or forgotten pseudoword. In order to increase the power, especially for conditions with less trials (e.g. forgotten word condition), we used data from both phases of Experiment 1 for the subsequent memory analyses. Since our analyses focused on lexical and memory differences rather than the manipulation of competitor priming, the potential confounding in phase two data should not be relevant here. Nevertheless, trial number was still unevenly distributed among these four conditions for each participant (see Table 4.1). Therefore, to minimise the bias of the mean gradiometer RMS for each condition due to the uneven number of trials, the RMS of each trial was calculated before robust averaging within each subject, though at the cost of lowering the signal to noise ratio. All trials were otherwise preprocessed following the same procedure as described in the section 2.2.6 of Chapter 2.

Sensor data from magnetometers and gradiometers were analysed separately. We converted the sensor data into 3D images (2D sensor x time) and performed F tests for main effects across sensors and time. Reported effects were obtained with a cluster-defining threshold of $p < .001$, and significant clusters identified as those whose extent (across space and time) survived $p < 0.05$ FWE-correction using Random Field Theory (Kilner & Friston, 2010). When plotting waveforms and topographies, data are shown for sensors nearest to the critical points in 2D image space. The main effects of interest were remembered vs forgotten words and remembered vs forgotten pseudowords.

In order to locate the brain sources of the effects found in sensor space, source reconstruction was also conducted in the same manner as described in section 2.2.7 of Chapter 2. The source inversion was projected onto the time window of 782-870ms post-DP based on the time span of the largest significant cluster from sensor space.

Table 4.1 The mean, maximal, minimal and standard deviation of the trial number of each condition in the subsequent memory analyses

Condition Name	Number of Trials			
	Mean	Max	Min	SD
Remembered Word	201	234	162	20.9
Forgotten Word	39	78	6	20.9
Remembered Pseudoword	141	163	118	12.4
Forgotten Pseudoword	99	122	77	12.4

4.3.2 Results & Discussion

We examined the main effect (sensor-time anova) of subsequent memory (remembered vs forgotten), interaction between memory and lexicality (words vs pseudowords), the difference between remembered and forgotten words, and the difference between remembered and forgotten pseudowords in magnetometers, gradiometers and source space. Figure 4.8A shows that the magnetometers detected a significant cluster (1313 sensor x time points, $p < .001$) in the right posterior region in which remembered items evoked more negative amplitude effects on the scalp than forgotten items at -114 to -73ms before the deviation point. Figure 4.8B shows

that this effect was driven more by the difference between remembered and forgotten words, $t(20) = -3.34$, $p = .006$ (Bonferroni corrected), in comparison to the difference between remembered and forgotten pseudowords, $t(20) = -2.66$, $p = .030$ (Bonferroni corrected). One possible reason for this is that the pre-DP segments in word items could already help differentiate word targets (e.g. *basin*) and semantic lures (e.g. *sink*), whereas this was unlikely for pseudoword targets (e.g. *basef*) and their phonological lures (e.g. *basoy*) that shared the same initial sounds. However, it was still possible that neural responses to remembered and forgotten pseudowords were different during the pre-DP period.

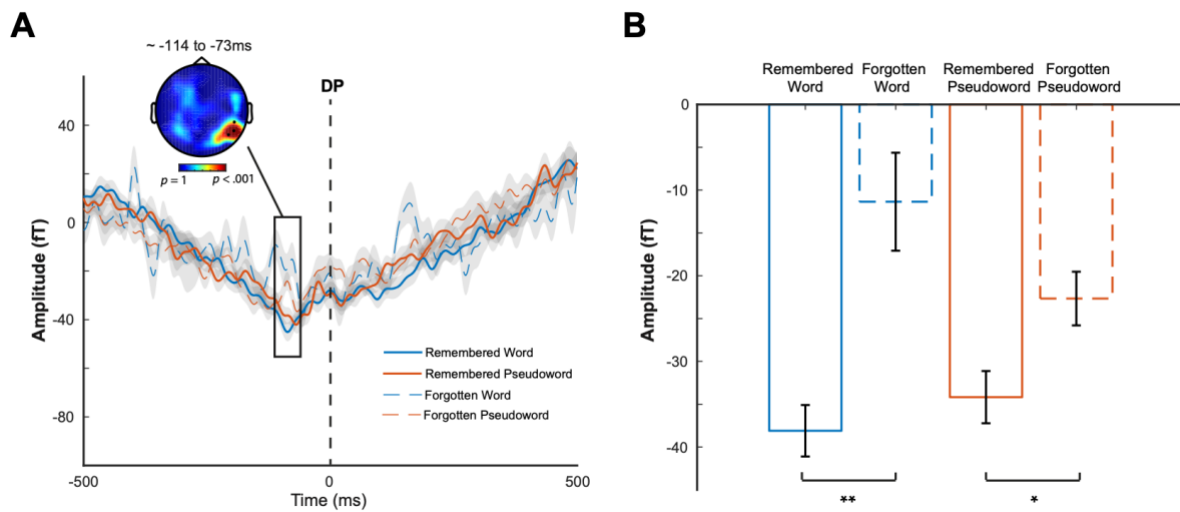


Figure 4.8. Cluster showing significant neural difference between remembered and forgotten items in Magnetometers. **A.** The topographic plot shows the statistically significant cluster. Waveforms represent magnetometer responses averaged over the spatial extent of the significant cluster shown in the topography. The grey shade of waveforms represents \pm within-participant SE, adjusted to remove between-participants variance. **B.** Signals evoked by conditions of interest averaged over temporal and spatial extent of the significant cluster shown in panel A. All error bars represent \pm within-participant SE, adjusted to remove between-participants variance. Statistical significance: ** $p < 0.01$, * $p < 0.05$.

The magnetometer sensors did not show any significant clusters for the interaction between memory and lexicality. The clusters for the difference between remembered and forgotten words did not survive $p < 0.05$ FWE-correction either. However, the magnetometers showed three significant clusters that reflected difference between remembered and forgotten pseudowords. Figure 4.9A shows the earliest cluster (1522 sensor x time points, $p < .001$) at 685-750ms post-DP in the right hemisphere in which more negative amplitude were shown for remembered pseudowords than forgotten pseudowords, while Figure 4.9B shows a larger cluster (2322 sensor x time points, $p < .001$) at 782-870ms post-DP in the left hemisphere in which remembered pseudowords evoked neural responses with more positive amplitude than forgotten ones. Given the proximity of time and location of the two clusters, it is likely that they came from the same source. Additionally, Figure 4.9C shows a cluster (1203 sensor x time points, $p < .001$) in the posterior area of the brain at 821-857ms post-DP in which remembered pseudowords showed more negative amplitude than forgotten ones. The gradiometer sensors did not show reliable effects in any of these contrasts, which could be due to the very uneven trial numbers in word conditions and low signal to noise ratio in gradiometers.

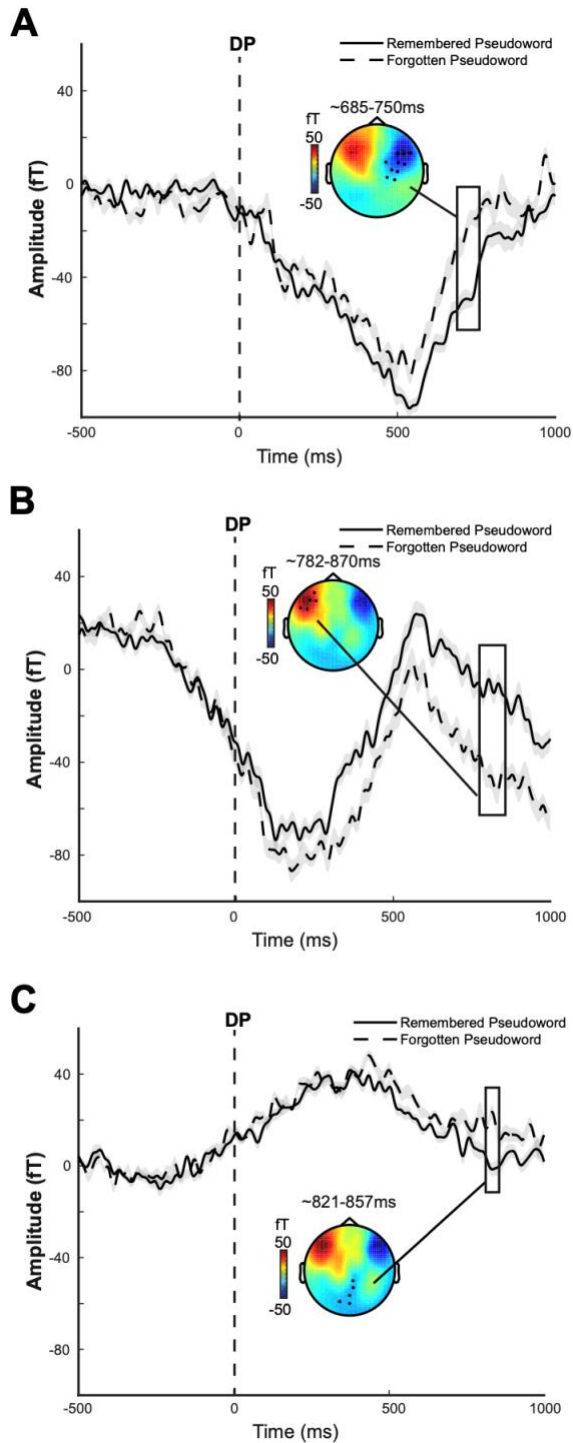


Figure 4.9. Clusters showing significant neural difference between remembered and forgotten pseudowords in Magnetometers over time. The topographic plots show the statistically significant clusters. Waveforms represent magnetometer responses averaged over the spatial extent of the significant cluster shown in the topography. The grey shade of waveforms represents \pm within-participant SE, adjusted to remove between-participants variance.

In source space (Figure 4.10), we localised pseudoword memory effects found in magnetometers over 782 to 870ms (the time span of the largest cluster) to the right lateral and medial temporal lobe (panel A, volume of 2555 voxels, $p < .05$, peak at $x = 62$, $y = -14$, $z = -22$, peak-level effect size $z = 2.71$) and the left anterior superior frontal gyrus (panel B, volume of 367 voxels, $p < .05$, peak at $x = -22$, $y = 52$, $z = 4$, peak-level effect size $z = 2.53$) with forgotten pseudowords evoking greater source strength than remembered pseudowords in these two clusters, $t(20) = -1.74$, $p = .097$ (panel D); $t(20) = -2.22$, $p = .038$ (panel E), and the left middle temporal gyrus (panel C, volume of 172 voxels, $p < .05$, peak at $x = -58$, $y = -46$, $z = -8$, peak-level effect size $z = 2.51$) with remembered pseudowords producing stronger responses than forgotten ones, $t(20) = 2.69$, $p = .014$ (panel F). Although these source localisation results were not significant at the cluster level, they reflected the source of the significant sensor time results, which are FWE cluster corrected. Since the canonical cortical meshes used for source reconstruction did not include the hippocampus, it is possible that the source shown in panel A and C could have originated from that region.

These observations seemed consistent with the ERP results from previous literature (Otten et al., 2008), in which a negative-going ERP was found for remembered versus forgotten pseudowords with a focus on right temporal sites at around 1000ms after stimuli onset. The engagement of the medial temporal lobe during this initial encoding of pseudowords is also in line with the complementary learning systems account. Moreover, pseudoword subsequent memory was also predicted by neural activity in the left anterior superior frontal gyrus. This could be due to cognitive control related to pseudoword identification, since the region is anatomically connected with the cingulate cortices (Li et al., 2013). Or it could be caused by

phonological processing during pseudoword encoding, as stronger activation in the nearby left inferior frontal gyrus has been observed for processing of phonological information of pseudowords relative to familiar words (Newman & Twieg, 2001; Poldrack et al., 1999) and that the magnitude of this activation should positively correlate with subsequent word memory (Clark & Wagner, 2003).

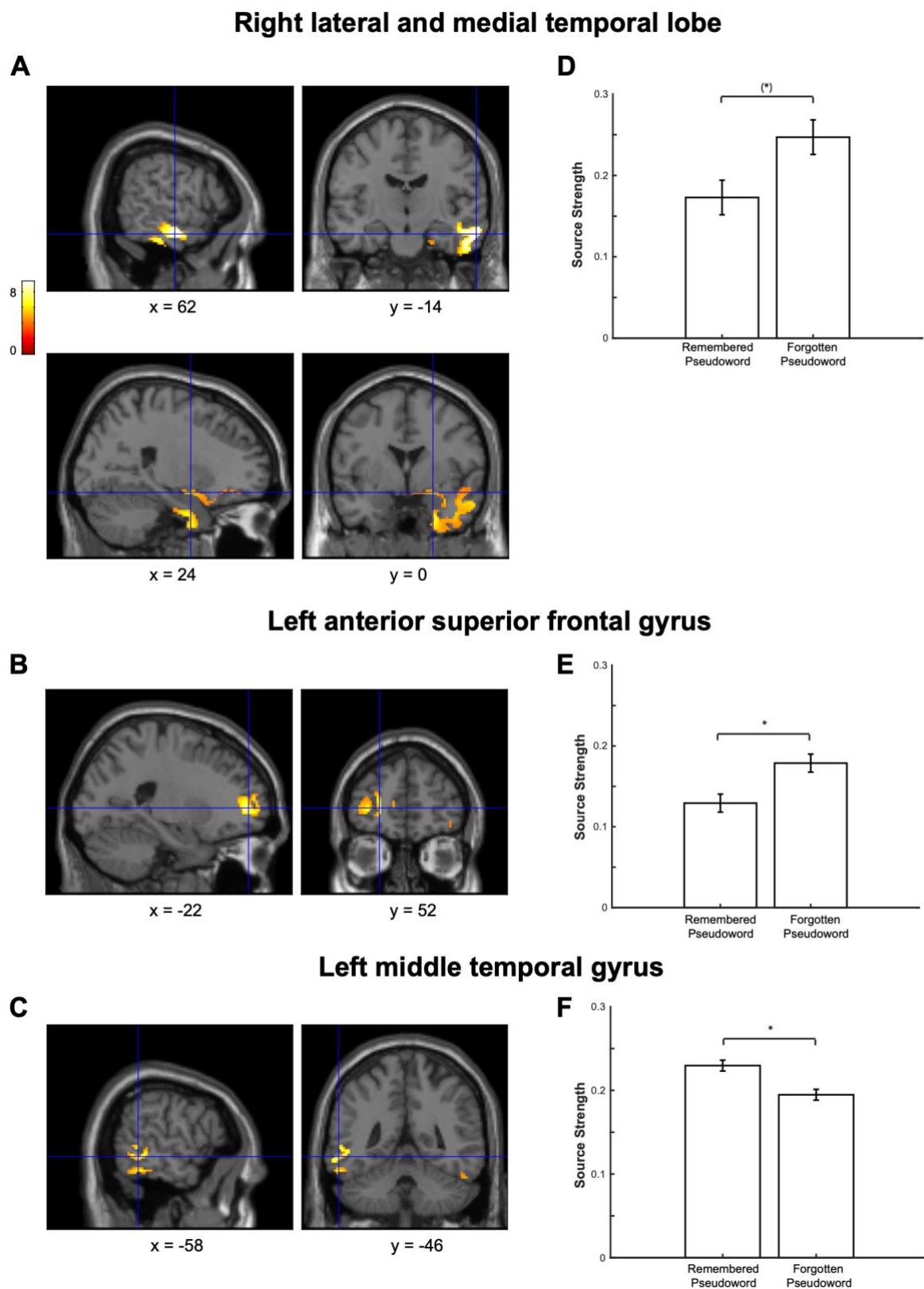


Figure 4.10. Neural activities predicting subsequent pseudoword memory localised in the source space. *A, B & C.* The difference of source activities to remembered and forgotten

pseudowords rendered onto the cross section of a single subject brain template, thresholded at FWE-corrected cluster-level $p < 0.05$. The x coordinate shows the sagittal view and the y coordinate shows the coronal view. *D, E & F*. The source strength evoked by remembered and forgotten pseudowords averaged over voxels within each cluster shown in panel A, B and C respectively. All error bars represent \pm within-participant SE, adjusted to remove between-participants variance. Statistical significance: * $p < 0.05$, (*) $p < 0.1$.

4.4 General discussion

This Chapter reported results from memory tasks of Experiment 1 to 5, which examined the effect of competitor priming on word and pseudoword episodic memory encoding; thereby testing whether priming-induced changes to computations of prediction error improved memory. At the same time, the use of lexical decision and pause detection tasks that respectively focused on lexical and sub-lexical processing during priming modulated the strength of prediction for target items.

We showed that, while the single prime paradigm in Experiment 1-3 did not evoke very clear or replicable effects potentially due to the lack of power, multiple prior presentations of the same competitor prime with a focus on lexical processing (Experiment 4b & 5b) evoked reliably better memory for target words than unprimed words, which is consistent with the PIMMS account (Henson & Gagnepain, 2010). The memory effects were not reliable for target words whose competitor primes were encoded using the pause detection task (Experiment 4a & 5a), which confirmed that lexical processing is at a deeper level than sub-lexical processing (Craik & Lockhart, 1972). Additionally, we did not find evidence for the “item-context trade-off” attentional resources account (Jurica & Shimamura, 1999), as shorter response times

during the encoding phase did not seem to have improved episodic memory of items encoded in the experimental context.

Finally, in terms of our subsequent memory analyses of the MEG data from Experiment 1, we did not find significant neural response supporting word encoding, possibly due to the uneven number of trials in word conditions. However, our results support that neural activity during pseudoword learning was localized to lateral and medial temporal lobe and left anterior superior frontal gyrus from ~782 to ~870ms post-DP. The locations and timing of these neural responses are consistent with previous literature (Clark & Wagner, 2003; Otten et al., 2007) and the initial rapid learning of novel words as supported by the CLS account (Davis & Gaskell, 2009).

4.4.1 Competitor priming effects modulate memory encoding

Chapter 2 and 3 reported studies that investigated the competitor priming effect during spoken word recognition, in which word-primed words evoked delayed behavioural responses and greater neural signals in the STG when participants' attention was on lexical processing. The results described in this chapter further examined the effect of competitor priming on word memory. We found that 1) stable effects of word priming that enhanced word memory were present only after multiple presentations of the competitor prime, 2) multiple presentations of pseudoword primes also evoked similarly positive effects on word memory as word primes, and 3) these effects were reliably observed only when participants' attention was on lexical processing but not phonological processing during the prime phase. I will discuss these findings as explained by different theories of memory encoding below.

The role of prediction error as proposed by the PIMMS account

Similar to the Predictive Coding account for spoken word recognition, the PIMMS account (Henson & Gagnepain, 2010) is also a Bayesian framework that suggests a hierarchical organization of human perception and memory and that the difference between higher-level predictions and lower-level signals, i.e. the prediction error, plays a key role in episodic memory encoding. Under the PIMMS framework, word priming provides a prior that enhances the prediction between the prime's initial and ending segments, therefore a target word that shares the same initial syllables but different ending should evoke a larger prediction error, which improves memory encoding. However, when the target or the prime is a pseudoword without pre-existing lexical and phonological representations, there is high uncertainty in the semantic memory system, which leads to less effective interactions with the episodic memory system, hence memory encoding and recollection is weak. Note that this is different from the maximal prediction error evoked by pseudowords in the perceptual systems during spoken word recognition.

Our study results showed that both word and pseudoword priming evoked significantly better word memory only after the same word prime was presented multiple times and that the prime task was the lexical decision task. No such effect was found for target pseudowords. These results have multiple implications, first of all, the fact that multiple presentations of the same competitor prime enhanced word memory but single presentation did not produce the same effect indicated that the stronger and more accurate the prediction was, the better the memory performance became. Our Experiment 4b is a replication of Experiment 1 and 2 except that the same prime was presented four times rather than a single time and that the lag between prime and target items was also longer (see methods in Chapter 3). Yet the memory effects of competitor priming were found only when the prime was presented multiple times. Similar effects were also found by Greve and colleagues (2017), who showed that memory encoding

for scene-face pairing was improved when the erroneous prime (the same scene paired with a different face) was presented multiple times during training. These results are in line with the PIMMS account that the size of the prediction error and hence the success of memory encoding is positively associated with the degree of precision of the prior.

However, the PIMMS account also predicts that a single competitor word prime should also induce prediction error on the target word and enhance its memory encoding. One possible reason why we did not find this effect in Experiment 1 and 2 could be that the number of participants in these two experiments were too small and hence there might be a lack of power to detect this effect. Another possible explanation is that a single prime with a long lag (20-80 trials) was simply too weak to form a strong enough prediction to modulate memory encoding.

In addition, it should also be noted that multiple presentations of the pseudoword prime also had a similar effect in improving word memory as word competitor primes (Experiment 4b and 5b), while an opposite negative effect was detected when there was only a single pseudoword prime (Experiment 1-3). The PIMMS account only predicts weak encoding for single pseudoword-primed word condition, but it does not predict any negative memory effect either. The disruptive effects shown in the earlier (potentially underpowered) experiments could be due to proactive interference from the prime item during memory retrieval process (Yonelinas, 2019). However, when the same pseudoword was presented multiple times, it might have activated similar sounding words. Hence, a stronger lexical prediction might have been formed, which evoked a similarly large prediction error as word primes.

In contrast to target words, there was no reliable memory effect for target pseudowords. As mentioned above, strong memory encoding and recollection relies on pre-existing phonological and semantic representations to a large extent, hence effective pseudoword learning may need many more learning episodes or long-term overnight consolidation (Davis

& Gaskell, 2009; James et al., 2017). Another reason could be that the pseudoword memory task was a phonological task that relied only on the ending segments to differentiate the target from the foil, hence memory trace of the initial stem could not facilitate the task performance. Therefore, the task might be too difficult to be sensitive to the experimental manipulation.

Levels of processing and attentional resources

Another important finding is that memory encoding effects were reliable only when participants focused on lexical processing (Experiment 4b and 5b) rather than phonological processing (Experiment 4a and 5a) during the prime phase. Indeed, according to the levels of processing model (Craik & Lockhart, 1972), lexical semantic processing is at a “deeper” level than phonemic processing, which should hence lead to better memory recall. It is possible that lexical processing enhanced predictions between the initial and the final segments of the prime, since the lexical decision task required participants to treat each item as a complete lexical-semantic unit. Therefore, larger prediction error was generated when perceiving the target, since its final segment violated the prior prediction. In contrast, during the pause detection task, participants mainly focused on lower-level phonetic features of prime items, and they did not have to treat each item as a complete lexical-semantic unit. In addition, each prime was also presented twice with pause inserted, which could disturb the prediction between pre- and post-DP segments.

On the other hand, our study results did not show evidence for the attentional resources account that proposes “item-context trade-off” (Jurica & Shimamura, 1999). The longer response times evoked by word-primed words in the encoding phase of Experiment 4b (see results in Chapter 3) indicated greater attentional resources spent processing the item. The fact that these words were also better remembered could not reflect any “trade-off” between item

and context memory. Moreover, response times were generally shorter when the encoding task was pause detection than when the task was lexical decision, but the memory accuracy was generally lower when pause detection was used during encoding. Hence, even if there was better episodic association between items and the experimental context during the pause detection task, it was not reflected in the memory results. These results were also consistent with a previous study by Kim and colleagues (2012), who found that reduced attention on tasks caused by prior item repetition also decreased episodic encoding of the item and context association rather than improving it. However, it should also be noted that we did not manipulate the experimental context by presenting the item using a specific source, such as different voices, or presenting the item together with another context stimuli, such as a background noise or picture, hence the experimental context varies depending on each participant's environment and mental status when they were responding to each stimuli, and we could not test participants' memory for the context explicitly. Therefore, our study design might not be optimal for testing the "trade-off" between item and context.

4.4.2 The neural mechanisms that support pseudoword encoding

Although we did not find reliable behavioural effects reflecting pseudoword learning, it did not mean that pseudowords were not encoded at all. The subsequent memory analyses of the MEG data allowed us to examine the neural underpinnings of word and pseudoword encoding. We did not find reliable clusters that represented word memory, possibly due to the very uneven number of trials in the remembered and forgotten word conditions. However, we found that pseudoword learning was driven by positive-going amplitudes in the left middle temporal gyrus and negative-going amplitudes in the right lateral and medial temporal lobe and the left anterior superior frontal gyrus between ~782 and ~870ms after the deviation point of stimuli.

The neural responses shown in the medial temporal lobe are in line with the initial rapid learning of novel words, as proposed by the complementary learning systems account (CLS; Davis & Gaskell, 2009), and consistent with the fMRI literature suggesting hippocampal contributions to word learning in general (Breitenstein et al., 2015; Davis et al., 2009; Mestres-Misse et al., 2008). The negative-going amplitudes shown in the right lateral and medial temporal lobe around 800ms post-DP were also in accordance with previous results found by Otten and colleagues (2008), in which pseudoword encoding was driven by negative-going ERPs at central and right temporal sites from 1000ms onwards post stimuli onset. Furthermore, neural activity in the left anterior frontal gyrus could be related to cognitive control for encoding, which has been found in the anatomically connected cingulate cortices (Li et al., 2013) and phonological processing of pseudowords, which has been found in the nearby left inferior prefrontal cortex and frontal gyrus (Clark & Wagner, 2003; Poldrack et al., 1999).

Due to the uneven number of trials in the word memory conditions, our results did not show a difference of neural underpinnings between word and pseudoword memory encoding. MEG might also not be the best method for measuring memory results, as the signals could be noisy and there was no hippocampal mesh for the forward model. However, the findings indicated that, despite a lack of significant behavioural pseudoword memory effect in the recognition memory task, neural responses in the medial temporal lobe still supported pseudoword encoding, which indicated initial rapid learning through the episodic systems.

4.4.3 Conclusion

Overall, this chapter reported the results of memory tasks from a series of studies using the competitor priming paradigm. We found that multiple presentations of the same word or pseudoword prime enhanced memory recollection of similar sounding target words when

participants focused on lexical processing during the prime task. These results are in line with the PIMMS account, which proposes that greater prediction error improves episodic memory. The neural results from the subsequent memory analyses of the MEG data also provided evidence that pseudoword encoding was driven by initial rapid learning in the medial temporal lobe, consistent with the literature that showed hippocampal contributions to novel word learning and the initial stage of the complementary learning systems.

Together with the findings reported in Chapter 2 and 3, computations of prediction error appear crucial in explaining both spoken word recognition and learning, which I will discuss further in the next chapter.

5 CONCLUDING REMARKS

The work reported in this thesis shows that computations of prediction error account for spoken word recognition and learning and that these processes are better supported by lexically informed processing compared to sub-lexical processing. This final chapter brings together the research reported in Chapter 2, 3 and 4 by summarizing and integrating the key findings. The limitations of these studies and future directions for research on spoken word recognition and learning are also discussed.

5.1 Summary

Chapter 2

The MEG study described in Chapter 2 tested whether the neural activities during spoken word recognition support the competitive-selection account (e.g. TRACE; McClelland & Elman, 1986) or predictive-selection account (e.g. Predictive Coding, Davis & Sohoglu, 2020). Through competitor priming, the prior probability of the activated lexical units was modulated differently for pre-DP and post-DP segments of the target items. While competitive-selection models and predictive-selection models both predict that word primed by a similar sounding word sharing the same initial segments are recognised more slowly, i.e. the competitor priming effect, they make different hypotheses for the timing and location of neural effects triggered by competitor priming. Our behavioural results showed slower response times to competitor primed words compared to unprimed words. Neural evidence recorded by MEG showed a lexical influence on neural activity localized to the STG, with competitively primed words showing stronger neural responses than unprimed words *after* the point at which they could be

uniquely identified but not *before*. Crucially, the elevated neural responses also correlated with the delayed recognition of competitor primed words.

The location, direction and timing of the neural activities recorded in the study were more consistent with the predictive-selection account. The links between behavioural and neural activities provided further evidence that prediction error plays a crucial role during spoken word recognition. It is true that studies using continuous speech (Brodbeck et al., 2018; Donhauser & Baillet, 2019) found effects accounting for both lexical entropy (which quantifies lexical competition) and surprisal (similar to prediction error), but these measures are highly correlated with each other in natural speech (Gwilliams & Davis, 2021) and it was hard to dissociate them in these studies.

Dufour and Nguyen (2017) found that the behavioural competitor priming effect was not modulated by voice change between the prime and the target when the prime was presented just once. However, when the prime was presented in a different voice multiple times, the competitor priming effect was attenuated. Future studies could explore whether the location and timing of corresponding neural effects were also affected by different episodic trace such as the change of voice. Moreover, manipulations of such episodic trace may also help dissociate the different processes involved in perceptual priming and recognition memory. Additionally, the perception of pseudowords in the current paradigm was limited to form, i.e. phonological level. It is worth further investigating whether pseudowords perceived with arbitrary semantic meaning would alter the neural implementations of the competitor priming effect.

Chapter 3

The behavioural experiments reported in Chapter 3 examined lexical and sub-lexical processing during spoken word recognition using the competitor priming paradigm. In

Experiment 4a, 4b, 5a and 5b, the tasks used during the prime phase aimed to dissociate the inhibitory lexical competition effect, caused by the prior presentation of competitor prime, and the facilitatory phonotactic priming effect, caused by the priming of ambiguous phonological segments when pause was present in the prime stimuli during the pause detection task. The tasks used in the target phase tested whether the probability change modulated by competitor priming could be affected by task-evoked decision (see Balota & Chumbley, 1984).

The findings showed that only when participants' attention was on lexical processing (lexical decision) during the perception of both prime and target items, the competitor priming effect (i.e. delayed recognition of primed words) was significant. However, when phonological task (pause detection) was used on both prime and target items and that word items were presented multiple times both with and without pause, there was a phonological facilitatory effect on target word recognition due to the activation of multiple compatible lexical items during the prime phase, even though the segments after pause were not consistent between prime and target word items. Conversely, when lexical decision and pause detection were mixed as the tasks for the perception of prime and target words, no reliable effects were found on word targets although there was a trend for slower response. Results from combined studies showed interaction effects between task and prime type for word targets, which differentiated the experiment using phonological task throughout both the prime and target phases (Experiment 5a) from the other experiments. Additionally, reliable facilitatory effects were also shown for pseudoword targets during pause detection when they were primed by the same pseudoword multiple times using the lexical decision task.

These findings showed opposite effects of lexical and phonological processing as suggested by the dual-level account (Vitevitch and Luce, 1998). Such effects from lexical and phoneme levels are in principle supported by localist models of spoken word recognition such as TRACE, but the facilitative phonological processing only affects the lexical level through

interaction and does not drive word recognition directly. In contrast, the Predictive Coding account incorporates both inhibitory lexical processing and facilitatory phonological processing, as calculations of lexically informed prediction error happen at the phoneme level. In addition, these opposite effects are also supported by the DCM model (Gaskell & Marslen-Wilson, 1997). This is because the DCM incorporates blended lexical and phonological representations in the same nodes that support both inhibitory lexical competitions and facilitatory phonotactic priming of coherent sound segments.

Additionally, the significant interaction effects between target task and prime type that differentiated Experiment 5a (which used the pause detection task in both phases) from 4a, 4b and 5b (which all used the lexical decision task at some point during the study) also indicated that task-evoked decision may affect spoken word recognition. While the lexical decision task requires knowledge of both word meaning and frequency (Balota & Chumbley, 1984), lexical representations of the stimuli may not even be accessed when the task (e.g. pause detection) did not require such processing. Furthermore, the pseudoword facilitatory priming effect on pseudoword targets (in Experiment 5b) could be because repeated lexical decision on the prime pseudowords activated multiple lexical items beginning with the same initial sounds. Therefore, the primed pre-DP segments had greater phonotactic probability, which led to easier perception of target pseudowords during the pause detection task.

I should also acknowledge that the lag between the prime and target items in the separate phase design (Experiment 4-5) might have been too long, which could explain why the effect of competitor priming was attenuated even when the lexical decision task was used for both the prime and target phases and that the prime was repeated for multiple times (Experiment 4b). This could also be the reason why no reliable difference was found between the results of Experiment 4b and the results of experiments that mixed lexical and sub-lexical processing in prime and target phases (Experiment 4a and 5b). On the other hand, the findings

also indicated that the facilitatory effect of phonological priming seemed to be robust and could last for ~7min on average when the prime was presented four times. Future studies could test whether such long-lasting effect is replicable in real life phoneme sequence learning using natural speech.

Chapter 4

Chapter 4 reported subsequent memory tasks on the items perceived during the MEG and behavioural experiments. These tasks investigated the effect of competitor priming on memory encoding. The competitor priming paradigm, together with lexical and sub-lexical processing manipulated by tasks during the encoding phase, modulated the strength and accuracy of predictions for target words. Therefore, the effect of prediction on subsequent memory of words could be assessed. Additionally, the neural activity associated with memory encoding was also examined through subsequent analyses of MEG signals. The behavioural results showed reliable encoding of competitor primed words when the prime was presented multiple times using the lexical decision task but not the pause detection task. There was no significant behavioural effect of competitor priming on memory for pseudoword targets. However, the neural activity associated with pseudoword memory encoding more generally was localized to the lateral and medial temporal lobe and the left anterior superior frontal cortex.

The behavioural results indicate that enhanced prediction error caused by competitor priming facilitated memory encoding of words when the encoding was repeated multiple times and involved deeper lexical-semantic processing. These findings are consistent with the PIMMS account (Henson & Gagnepain, 2010) and the levels of processing model (Craik & Lockhart, 1972). They are also related to the results shown in Chapter 3, which indicated different effects of lexical and sub-lexical processing on word perception. The neural

implementations of pseudoword encoding localized to the medial temporal lobe was supported by lesion studies (Bayley et al., 2008; Martins et al., 2006) and functional neural imaging evidence (Breitenstein et al., 2005; Davis et al., 2009), and was consistent with the CLS account (Davis & Gaskell, 2009).

It should be noted that the different memory tasks used for words (with semantic foil) and pseudowords (with phonological foil) might have made it more difficult to retrieve pseudoword targets than word targets. The fact the words have prior semantic representations while pseudowords do not also made it harder to encode and recollect pseudowords. Therefore, future studies could reduce the difficulty of pseudoword encoding by adding more repetitions, reducing the lag between encoding and retrieval tasks or presenting pseudowords with associable semantic cues. Additionally, while these experiments only explored effects of competitor priming on the memory of newly encountered pseudowords, it is worth exploring whether novel items after overnight consolidation (hence lexicalized) can show stronger episodic memory encoding after being primed by competitor words.

5.2 Predictive coding for spoken word recognition and learning

The studies reported in this thesis showed that computations of prediction error could explain both spoken word recognition and memory encoding, although the predictions required for the two processes may be different.

For spoken word recognition, lexically informed top-down predictions are updated by bottom-up prediction errors iteratively during speech perception (Davis & Sohoglu, 2020). Results from Chapter 2 and 3 showed that one presentation of the competitor word was sufficient to produce slowed behavioural and increased neural responses. This is explained by the large prediction error evoked during the perception of the target word. However, a long lag

between prime and target could also attenuate this change in prediction error. Conversely, for words to be effectively encoded, it is necessary that there is sufficient synaptic exchange (learning) between the higher prediction level and lower sensory level (Henson & Gagnepain, 2010). The results presented in Chapter 4 indicated that the strength of prediction, i.e. multiple presentations of the same competitor word, was crucial, while the delay between prime and target items did not seem to matter as much as during word perception. Importantly, both processes require lexical-level processing. The fact that pseudowords do not have prior lexical representations made it hard for their perception and learning to be affected by the manipulation of prediction.

Previous studies (Church & Schacter, 1994; Schacter et al., 1993) have found dissociations between implicit recognition memory, which was shown by perceptual priming in the absence of conscious recollection, and explicit recognition, which was tested by deliberate memory tasks. While spoken word priming was shown to be affected by voice, intonation and fundamental frequency, these factors had no effects on explicit recognition task (Church & Shacter, 1994; See Dufour & Nguyen, 2017 for evidence on competitor priming). Similarly, Schacter and colleagues (1993) reported that patients with amnesia exhibited normal priming effects on degraded word items but performed poorly during the explicit recognition memory task.

This dissociation was possibly because conceptually driven process was more engaged in explicit memory retrieval compared to word perception (and perhaps also due to the use of semantic foil in our word memory task). Hence, multiple presentations of the same word item might have particularly strengthened lexical or semantic-based predictions for the prime word, hence prediction error evoked by the target word was larger, which improved its memory encoding and retrieval. Whereas delayed spoken word recognition caused by competitor priming was more due to pre-lexical segment-level perceptual mismatch caused by lexically

informed predictions for the competitor word (as shown in the neural evidence from Chapter 2), which might be more sensitive to the delay between prime and target items. Such difference between recognition memory and perceptual priming was also reflected in pseudowords. While maximal perceptual prediction error was evoked by pseudowords during spoken word recognition, the prediction error that drives interactions between the episodic and semantic memory systems was small due to their lack of pre-existing representations. However, it is true that both the effect of delayed spoken word recognition and improved memory are associated with lexical processing, which I will discuss in more detail in the next section.

5.3 Lexical and sub-lexical processing for spoken word recognition and learning

It has been recognised by most spoken word recognition models (DCM, Predictive Coding, Shortlist, TRACE) that the identification of spoken word involves at least two levels of processing, lexical and sub-lexical. However, the effects of the two processes seem to be in opposite directions, with high neighbourhood density producing inhibitory lexical-level competitions and high phonotactic probabilities facilitating identification of sub-lexical sound sequence (Gaskell & Marslen-Wilson, 2002; Vitevitch and Luce, 1998), as reflected by the delayed response caused by competitor priming and faster response caused by phonological priming of pre-DP segments described in Chapter 3. Similarly, as shown in Chapter 4, word memory is also affected differently by lexical and sub-lexical level processing, particularly because lexical level processing tested through the lexical decision task is a deeper form of processing with greater semantic involvement, while sub-lexical level processing tested using the pause detection task is a shallower form of encoding (Craik & Lockhart, 1972; Craik & Tulving, 1975).

However, the perception or learning of pseudowords are less clearly affected by these processing. The only reliable effect on pseudoword perception found in our studies was the facilitatory effect when a pseudoword was primed by a similar sounding pseudoword multiple times. This could be explained by associative learning between lexical responses and the initial segments of items (Experiment 4b) and phonotactic facilitation caused by matching of the speech sounds shared by multiple activated lexical representations (Experiment 5b). However, no competitor priming effect was shown for pseudoword targets due to maximal prediction error produced during their perception. Similarly, we did not find any reliable effects of pseudoword learning regardless of lexical or sub-lexical level processing during the encoding phase. Again, the fact that pseudowords do not have any pre-existing lexical representations may be the reason for these results. While sub-lexical facilitatory effects were still shown for pseudoword perception because of their shared pre-DP segments with words, lexical level inhibitory effects and semantic-related encoding process did not apply to them. Future studies on pseudoword perception and learning could consider involving lexical knowledge by providing semantic context during the perception of pseudowords or lexicalizing pseudowords in advance through training with arbitrary semantic cues (Takashima et al., 2014, 2017) or overnight consolidations (Davis & Gaskell, 2009). These results also suggested that, for language learning in real life, associate novel words with existing semantic knowledge or schema could be especially helpful for learning.

5.4 Conclusion

In this thesis, I reported studies that examined neural and functional mechanisms of spoken word recognition and learning using the competitor priming paradigm. These studies provided strong evidence for a unified account that computations of prediction error drive spoken word

recognition and memory encoding while the effects of lexical and sub-lexical processing were dissociated during these two processes. This research uniquely contributes to existing literature by bringing together spoken word recognition and learning with a common prediction-based framework using both neural and behavioural approaches. It also paves the way towards a mechanistic understanding of the two common human cognitive functions that are closely related to each other. Finally, it provides a solid base and a starting point for future research to further examine the role of prediction error and the relationship between lexical and sub-lexical processing in spoken word recognition and memory encoding.

APPENDIX A List of Stimuli

PhonDISC transcription (which uses unsyllabified, DISC character set) is provided for each stimulus.

1. Stimuli words and pseudowords

Prime	PhonDISC	Target	PhonDISC	Prime	PhonDISC	Target	PhonDISC
Words		Words		Pseudowords		Pseudowords	
letter	lEt@R	lettuce	lEtIs	letto	lEt5	lettan	lEt{n
boycott	b4kQt	buoyant	b4@nt	boyten	b4t@n	boymid	b4mId
canister	k{nIst@R	cannibal	k{nIb@l	canniphyll	k{nIfIl	cannittock	k{nIt@k
tribute	trIbjut	tribune	trIbjun	tribuke	trIbjuk	tribule	trIbjum
amid	@mId	amiss	@mIs	amic	@mIk	amip	@mIp
poker	p5k@R	pokey	p5kI	pokoa	p5k5	pokra	p5kr@
shallow	S{l5	chalet	S{lI	challon	S{l@n	shallume	S{ljum
hammer	h{m@R	hammock	h{m@k	hamment	h{m@nt	hamel	h{m@l
motor	m5t@R	motive	m5tIv	moton	m5tQn	motay	m5tI
kilo	kiI5	kiwi	kiwi	kini	kini	keebo	kib5
border	b\$d@R	boredom	b\$d@m	bordle	b\$d@l	bordate	b\$d@t
fillet	fIlIt	filly	fIli	fillow	fIl5	fillain	fIl@n
torrent	tQr@nt	torrid	tQrId	torrupt	tQrVpt	torroon	tQrun
minus	m2n@s	minor	m2n@R	minum	m2n@m	minarch	m2n@k
deacon	dik@n	decoy	dik4	deakime	dik2m	deaket	dikIt
crocodile	krQk@d2l	crockery	krQk@rI	crockellent	krQk@l@nt	crocony	krQk@ni

Appendix A

venue	vEnju	venom	vEn@m	venid	vEnId	venaut	vEn\$t
hijack	h2_{k	hygiene	h2_in	higent	h2_@nt	hijure	<u>h2_jU@</u>
carnival	k#nIv@l	carnivore	k#nIv\$R	carnivit	k#nIvIt	carnivest	k#nIvEst
climate	kl2mIt	climax	kl2m{ks	climod	kl2mQd	climain	kl2m1n
humour	hjum@R	humid	hjumId	humoy	hjum4	humite	hjum2t
burden	b3d@n	birdie	b3di	birdack	b3d{k	birdict	b3dIkt
clutter	klVt@R	cluster	klVst@R	clunnel	klVn@l	cluddon	klVd@n
counter	k6nt@R	counsel	k6ns@l	coundial	k6ndI@l	counger	k6n_@R
embassy	Emb@sI	embryo	EmbrI5	embicate	EmbIk@t	embuty	Embjuti
cemetery	sEmItrI	seminar	sEmIn#R	semidure	sEmIdju@	semipon	sEmIpQn
docile	d5s2l	dosage	d5sI_	doeset	d5sEt	doesop	d5s@p
crumpet	krVmpIt	crumple	krVmp@l	crumpoon	krVmpun	crumpon	krVmpQn
adverse	{dv3s	adverb	{dv3b	advertch	{dv3J	adverk	{dv3k
foliage	f5lII_	folio	f5lI5	folimon	f5lImQn	foliland	f5lII@nd
title	t2t@l	tighten	t2t@n	titum	t2t@m	titac	t2t@k
pilot	p2l@t	pylon	p2l@n	pielage	p2l@_	pilal	p2l@l
beaker	bik@R	beacon	bik@n	beacal	bik@l	beacus	bik@s
barrel	b{r@l	barren	b{r@n	barrod	b{r@d	barroph	b{r@f
cuckoo	kUku	cookie	kUKI	coogle	kUk@l	cookuse	kUkjuz
forage	fQrI_	foray	fQr1	forour	fQr@	forrack	fQr{k
summit	sVmIt	summon	sVm@n	summoy	sVm4	summack	sVm{k
banquet	b{NkwIt	bankrupt	b{NkrVpt	bankume	b{Nkjum	bankip	b{NkIp
ally	{l2	alloy	{l4	allop	{l@p	allent	{l@nt

Appendix A

village	vIII_	villain	vII@n	villief	vIIif	villate	vII1t
vertigo	v3tIg5	vertebra	v3tIbr@	vertifix	v3tIfIks	vertimid	v3tImId
uniform	junIf\$m	unicorn	junIk\$n	unistall	junIst@l	unigree	junIgri
gusty	gVsti	gusto	gVst5	gustom	gVst@m	gustide	gVst2d
turkey	t3kI	turquoise	t3kw4z	turkoon	t3kun	turkus	t3k@s
daily	d1II	dainty	d1ntI	daipent	d1p@nt	daimous	d1m@s
robot	r5bQt	rodent	r5d@nt	roculus	r5k@s	roton	r5S@n
garbage	g#bI_	garble	g#b@l	garbet	g#b1	garbide	g#b2d
dirty	d3tI	dervish	d3vIS	dircle	d3k@l	durnace	d3nIs
paradise	p{r@d2s	paradigm	p{r@d2m	paradite	p{r@d2t	paradike	p{r@d2k
succeed	s@ksid	succinct	s@ksINkt	succsence	s@ks@ns	succside	s@ks2d
furnace	f3nIs	furnish	f3nIS	furnet	f3nIt	furnic	f3nIk
mercy	m3sl	murmur	m3m@R	mertus	m3t@s	merbute	m3bjut
granite	gr{nIt	granule	gr{njul	granarch	gr{n#k	granult	gr{n@lt
ticket	tIkIt	tickle	tIk@l	tickod	tIkQd	tickute	tIkjut
voter	v5t@R	vocab	v5k{b	voble	v5b@l	vosey	v5zi
fashion	f{S@n	fascist	f{SIst	fashew	f{Su	fassure	f{SU@
wagon	w{g@n	waggle	w{g@l	waggus	w{g@s	wagget	w{g@t
bonnet	bQnIt	bonny	bQni	bonough	bQnVf	bonnome	bQn5m
domino	dQmIn5	domicile	dQmIs2l	dommiree	dQmIri	dommippet	dQmIpIt
credit	krEdIt	crevice	krEvIs	cremble	krEmb@l	creckon	krEk@n
gallery	g{l@rI	galaxy	g{l@ksI	gallagee	g{l@_i	gallaby	g{l@bi
ferry	fEri	ferret	fErIt	ferrer	fEr@	ferrack	fEr{k

Appendix A

crystal	krIst@l	christen	krIs@n	chrisic	krIsIk	chrisire	krIs2@
basis	b1sIs	basin	b1s@n	basef	b1sef	basoy	b1s4
pewter	pjut@R	putrid	pjutrId	pewtay	pjutl	pewtarb	pjut#b
funny	fVnI	funnel	fVn@l	funarm	fVn#m	funnane	fVn1n
envy	Envi	envoy	Env4	envice	Env2s	envam	Env{m
salad	s{l@d	sallow	s{l5	sallack	s{l{k	salius	s{l7s
insulin	InsjUlIn	insulate	InsjUl1t	insuloon	InsjUlun	insullock	InsjUlQk
diverse	d2v3s	divulge	d2vVl_	divite	d2v2t	divom	d2v@m
violent	v2@l@nt	violet	v2@l@t	viollus	v2@l@s	violumn	v2@l@m
fluid	fluId	fluent	flu@nt	fluack	flu{k	fluance	flu#ns
cyclist	s2klIst	psychic	s2kIk	cyker	s2k@R	cycrite	s2krIt
river	rIv@R	rivet	rIvIt	rivey	rIvi	rivest	rIvEst
bowler	b5l@R	boulder	b5ld@R	bowllion	b5l7n	bowlcha	b5lJ@
malice	m{lIs	mallet	m{lIt	malliff	m{lIf	mallin	m{lIn
tender	tEnd@R	tendon	tEnd@n	tendal	tEnd@l	tendus	tEnd@s
solar	s5l@R	solo	s5l5	solim	s5lIm	solay	s5l1
bucket	bVklIt	buckle	bVk@l	buckoon	bVkun	buckorde	bVk\$d
sequel	sikw@l	sequin	sikwIn	sequoy	sikw4	sequash	sikwQS
general	_En@r@l	generate	_En@r1t	generot	_En@rQt	generiff	_En@rIf
dragon	dr{g@n	drastic	dr{stIk	draffle	dr{f@l	dralot	dr{lQt
forever	f@rEv@R	forensic	f@rEnsIk	forredom	f@rEd@m	forremid	f@rEmId
cherry	JEri	cherish	JErIS	cheret	Jerl	cherald	JEr@ld
final	f2n@l	finite	f2n2t	finage	f2nI_	finect	f2nEkt

Appendix A

customer	kVst@m@R	custody	kVst@dI	custom	kVst@n@m	custopus	kVst@p@s
parcel	p#s@l	parson	p#s@n	parsus	p#s@s	parssad	p#s@d
local	l5k@l	locust	l5k@st	loakon	l5k@n	loaker	l5k@R
kidney	kIdnI	kidnap	kIdn{p	kidnate	kIdn1t	kidnise	kIdn2z
founder	f6nd@R	fountain	f6ntIn	founshion	f6nS@n	foundger	f6n_@R
matron	m1tr@n	matrix	m1trIks	matrod	m1trQd	matrope	m1tr5p
doodle	dud@l	duvet	duv1	doocon	duk@n	ducid	dusId
lion	l2@n	liar	l2@R	lieage	l2@_	lieette	l2@t
hazel	h1z@l	hazy	h1zI	hazar	h1z#	hazoon	h1zun
loosen	lus@n	lucid	lusId	loosack	lus{k	loosire	lus2@
vulgar	vVlg@R	vulture	vVIJ@R	vulgeon	vVl_@n	vuldom	vVld@m
mortar	m\$t@R	mortal	m\$t@l	mortus	m\$t@s	mortac	m\$t@c
female	fim1l	femur	fim@R	feamy	fimi	femote	fim5t
pinnacle	pIn@k@l	pinafore	pIn@f\$R	pinallor	pIn@l@R	pinadit	pIn@dIt
former	f\$m@R	format	f\$m{t	formish	f\$mIS	formaze	f\$m1z
career	k@r7R	caress	k@rEs	carute	k@rut	caroy	k@r4
franchise	fr{nJ2z	frantic	fr{ntIk	frander	fr{nd@R	franlot	fr{nIQt
slipper	sIip@R	sliver	sIiv@R	slidden	sId@n	slibbet	sIbEt
liver	lIv@R	livid	lIvId	livute	lIvjut	liverse	lIv3s
stubborn	stVb@n	stubble	stVb@l	stubbous	stVb@s	stubber	stVb@R
labour	l1b@R	label	l1b@l	labeth	l1b@T	labon	l1b@n
bullet	bUIIt	bullock	bUI@k	bullize	bUI2z	bullete	bUIIt
drizzle	drIz@l	drivel	drIv@l	driggle	drIg@l	drinom	drIn@m

Appendix A

table	t1b@l	taper	t1p@R	tamous	t1m@s	taget	t1gEt
damsel	d{mz@l	damson	d{mz@n	damser	d{mz@R	damsus	d{mz@s
delicate	dElIk@t	delegate	dElIg@t	deleral	dElIr@l	delepute	dElIpjut
soccer	sQk@R	socket	sQkIt	sockite	sQk2t	sockob	sQkQb
complement	kQmplIm@nt	complicate	kQmplIk1t	complinour	kQmplIn@R	complidence	kQmplId@ns
flourish	flVrIS	fluster	flVst@R	fluppent	flVp@nt	flummal	flVm@l
permanent	p3m@n@nt	permeate	p3mI1t	permuda	p3mjud@	permoiler	p3m4l@R
broker	br5k@R	brochure	br5S@R	broter	br5t@R	broget	br5gEt
tartan	t#t@n	tartar	t#t@R	tartal	t#t@l	tarttock	t#t@k
quarry	kwQri	quarrel	kwQr@l	quarrest	kwQrEst	quarid	kwQrId
lumpy	lVmpl	lumber	lVmb@R	lumket	lVmkIt	lumtice	lVmtis
merry	mEri	merit	mErIt	merrant	mEr@nt	merrope	mEr5p
ransom	r{ns@m	rancid	r{nsId	rancike	r{ns2k	ransorn	r{ns\$n
billion	blIj@n	billow	bII5	billid	bIIId	billous	bII@s
relish	rElIS	relic	rElIk	relip	rElIp	relid	rElId
journey	_3ni	journal	_3n@l	journume	_3njum	journet	_3nIt
plumber	plVm@R	plummet	plVmIt	plummock	plVm@k	plummute	plVmjut
trifle	tr2f@l	tripod	tr2pQd	trimate	tr2m1t	tricer	tr2s@R
cinema	sIn@m#	cinnamon	sIn@m@n	cinnamope	sIn@m5p	cinnamute	sIn@mjut
tacit	t{sIt	tassel	t{s@l	tassoy	t{s4	tassike	t{s2k
heritage	hErItI_	heresy	hErIsI	herelade	hErIl1d	herebin	hErIbIn
valley	v{li	valid	v{IId	vallous	v{l@s	vallar	v{l#
ruby	rubi	rhubarb	rub#b	rubon	rubQn	rubule	rubjul

Appendix A

total	t5t@l	totem	t5t@m	totus	t5t@s	tottock	t5t@k
transfer	tr{nsf3R	transfix	tr{nsfIks	transfike	tr{nsf2k	transfume	tr{nsfjum
teeny	tini	teeter	tit@R	teecon	tikQn	teebid	tibId
prodigy	prQdI_I	prodigal	prQdIg@l	prodicue	prQdIkju	proditus	prQdIt@s
volume	vQljum	volley	vQII	volance	vQl@ns	vollike	vQl2k
scandal	sk{nd@l	scanty	sk{ntI	scanpate	sk{np1t	scanvage	sk{nvI_
rhino	r2n5	rifle	r2f@l	rical	r2k5	ripot	r2pQt
mustard	mVst@d	mustang	mVst{N	mustine	mVst2n	mustic	mVstIk
cola	k5l@	colon	k5lQn	colit	k5lIt	colo	k5l5
rubbish	rVbIS	rubber	rVb@R	rubboa	rVb5	rubbut	rVbjut
mistress	mIstrIs	mystic	mIstIk	mistume	mIstjum	mistol	mIst@l
salute	s@lut	saloon	s@lun	saloupe	s@lup	saloosh	s@luS
brandy	br{ndi	brandish	br{ndIS	brandile	br{nd2l	brandos	br{ndQs
wicked	wIkId	wicket	wIkIt	wickic	wIkIk	wickiss	wIkIs
fatal	f1t@l	fable	f1b@l	facon	f1k@n	fajor	f1_@
super	sup@R	supine	sup2n	supave	sup1v	supid	supId
language	l{NgwI_	languish	l{NgwIS	languin	l{NgwIn	languil	l{NgwIl
series	s7riz	serum	s7r@m	serack	s7r{k	serot	s7rQt
modern	mQd@n	modest	mQdIst	moduce	mQdjus	modron	mQdr@n
swallow	swQl5	swaddle	swQd@l	swabute	swQbjut	swacket	swQkIt
accent	{ksEnt	axiom	{ks7m	accsort	{ks\$t	aksume	{ksjum
goblet	gQbIIt	goblin	gQbIIn	goblidge	gQbII_	goblice	gQbIIs
platinum	pl{tIn@m	platypus	pl{tIp@s	platymer	pl{tIm@R	platicate	pl{tIk@t

Appendix A

polish	pQlIS	pollen	pQl@n	pollack	pQl{k	pollar	pQl#
media	midj@	median	midj@n	meediage	midj@_	meediape	midj@p
foyer	f4l	foible	f4b@l	foitet	f4tEt	foikon	f4k@n
sherry	SEri	sheriff	SErIf	sherra	SEr@	sherrow	SEr5
spanish	sp{nIS	spaniel	sp{nj@l	spanode	sp{n5d	spanum	sp{n@m
pedigree	pEdIgri	pedestal	pEdIst@l	pedio	pEdI5	pedivore	pEdIv\$
protest	pr5tEst	proton	pr5tQn	protar	pr5t#	protal	pr5t@l
walnut	w\$lnVt	walrus	w\$lr@s	waldom	w\$ld@m	walver	w\$lv@
gable	g1b@l	geisha	g1S@	geivy	g1vi	geirret	g1rIt
witness	wItnIs	Whitsun	wIts@n	witcoy	wItk4	witmer	wItm@R
caper	k1p@R	kapok	k1pQk	capid	k1pId	capute	k1pjut
dowdy	d6dI	doughty	d6tI	dowcal	d6k@l	dowcid	d6sId
bunker	bVNk@R	bungle	bVNg@l	bunkton	bVNkt@n	bunktion	bVNkS@n
grubby	grVbI	grumble	grVmb@l	grugit	grVgIt	gruvice	grVvIs
sabre	s1b@R	sable	s1b@l	sabon	s1b@n	saboured	s1b@d

2. Target stimuli and foils

Word Stimuli	Word Foils	Pseudoword Stimuli	Pseudoword Foils
lettuce	cabbage	lettan	letty
buoyant	resilient	boymid	boymid
cannibal	savage	cannitock	cannidon
tribune	official	tribule	tribuse
amiss	mistake	amip	amidge
pokey	tiny	pokra	pokate
chalet	cabin	shallume	shallage
hammock	sedan	hamel	hammop
motive	reason	motay	motil
kiwi	grapefruit	keebo	keefit
boredom	tedium	bordate	bordus
filly	foal	fillain	filect
torrid	hot	torroon	torrium
minor	junior	minarch	minant
decoy	distract	deaket	deacal
crockery	cutlery	crocony	crocketal
venom	poison	venaut	venoid
hygiene	sanitation	hijure	hijous
carnivore	predator	carnivest	carnivoke
climax	apex	climain	climal

humid	moist	humite	humal
birdie	eagle	birdict	birdle
cluster	bunch	cluddon	clucket
counsel	advice	counger	counvas
embryo	zygote	embuty	embanent
seminar	assembly	semipon	semigel
dosage	measure	doesop	doesoll
crumple	crease	crumpon	crumpass
adverb	adjective	adverk	advern
folio	catalogue	foliland	foleor
tighten	constrict	titac	titage
pylon	tower	pilal	pieluc
beacon	signal	beacus	beacra
barren	infertile	barroph	barrip
cookie	biscuit	cookuse	cookrete
foray	raid	forrack	forroy
summon	convene	summack	summaut
bankrupt	insolvent	bankip	bankle
alloy	blend	allent	allade
villain	scoundrel	villate	villack
vertebra	spine	vertimid	vertilla
unicorn	phoenix	unigree	unipon
gusto	delight	gustide	gustick

turquoise	sapphire	turkus	turkon
dainty	elegant	daimous	daigel
rodent	mouse	roton	ropaque
garble	distort	garbide	garbon
dervish	devil	durnace	derpent
paradigm	pattern	paradike	paradire
succinct	concise	succside	succsorm
furnish	supply	furnic	furneal
murmur	whisper	merbute	merpet
granule	particle	granult	granol
tickle	stroke	tickute	tickom
vocab	jargon	vosey	vobot
fascist	racist	fassure	fasheen
waggle	jiggle	wagget	wagoy
bonny	pretty	bonnome	bonarch
domicile	residence	dommippet	dommitted
crevice	crack	creckon	crepit
galaxy	universe	gallaby	gallaty
ferret	rabbit	ferrack	ferrent
christen	baptize	chrisire	chrislot
basin	sink	basoy	basoy
putrid	rotten	pewtarb	pewteck
funnel	pipe	funnane	funnoob

envoy	emissary	envam	envout
sallow	pale	salius	salloy
insulate	isolate	insullock	insolute
divulge	reveal	divom	divoke
violet	lilac	violumn	violac
fluent	eloquent	fluance	flurette
psychic	telepath	cycrite	cycoy
rivet	pin	rivest	rivel
boulder	stone	bowlcha	bowlom
mallet	hammer	mallin	mallish
tendon	sinew	tendus	tendock
solo	single	solay	soloy
buckle	clasp	buckorde	buckuum
sequin	spangle	sequash	sequid
generate	produce	generiff	generen
drastic	severe	dralot	drapute
forensic	detective	forremid	forremel
cherish	treasure	cherald	cherrow
finite	limited	finect	finaut
custody	detention	custopus	custobal
parson	vicar	parssad	parser
locust	grasshopper	loaker	loakid
kidnap	abduct	kidnise	kidnus

fountain	cascade	foundger	founlot
matrix	vector	matrope	matrus
duvet	quilt	ducid	dootal
liar	fibber	lieette	lieus
hazy	misty	hazoon	hazent
lucid	clear	loosire	loosorn
vulture	buzzard	vuldom	vulneal
mortal	dying	mortac	mortent
femur	humerus	femote	femium
pinafore	apron	pinadit	pinamen
format	layout	formaze	formap
caress	cuddle	caroy	carid
frantic	panic	franlot	frangy
sliver	slice	slibbet	slinette
livid	angry	liverse	livoice
stubble	beard	stubber	stubbop
label	tag	labon	labute
bullock	cattle	bullete	bulleaze
drivel	nonsense	drinom	dridden
taper	narrow	taget	tasom
damson	plum	damsus	damsy
delegate	minister	delepute	delety
socket	plug	sockob	sockule

complicate	complex	complidence	compligy
fluster	agitate	flummal	fluzon
permeate	infuse	permoiler	permaigrate
brochure	catalogue	broget	bromel
tartar	plaque	tarttock	tartice
quarrel	squabble	quarid	quarren
lumber	timber	lumtice	lumdle
merit	virtue	merrope	merrod
rancid	fetid	ransorn	ransute
billow	wave	billous	biller
relic	artefact	relid	rellis
journal	diary	journet	journous
plummet	plunge	plummute	plumal
tripod	trivet	tricer	trylon
cinnamon	nutmeg	cinnamute	cinamy
tassel	bobble	tassike	tassix
heresy	sacrilege	herebin	herecker
valid	legal	vallar	vallow
rhubarb	pumpkin	rubule	rubet
totem	icon	tottock	tottad
transfix	stare	transfume	transfeat
teeter	wobble	teebid	teedom
prodigal	wasteful	proditus	prodelong

volley	strike	vollike	volloon
scanty	insufficient	scanvage	scanlin
rifle	shotgun	ripot	rhigette
mustang	stallion	mustic	musttle
colon	comma	colo	colite
rubber	elastic	rubbutte	rubyme
mystic	magic	mistol	mistard
saloon	tavern	saloosh	salule
brandish	wield	brandos	brandiot
wicket	cricket	wickiss	wickiff
fable	legend	fajor	failiff
supine	prostrate	supid	supoy
languish	decline	languil	languiff
serum	plasma	serot	serude
modest	humble	modron	moduct
swaddle	envelope	swacket	swatton
axiom	maxim	aksume	accsage
goblin	elf	gobllice	goblive
platypus	duck	platicate	plattivate
pollen	spore	pollar	polloy
median	middle	meediape	meediace
foible	fault	foikon	foyack
sheriff	sergeant	sherrow	sherrain

spaniel	poodle	spanum	spannard
pedestal	platform	pedivore	pedicap
proton	positron	protal	proty
walrus	sealion	walver	walgar
geisha	singer	geirret	geinham
Whitsun	Easter	witmer	witlen
kapok	wadding	capute	capent
doughty	tough	dowcid	dowlet
bungle	botch	bunktion	buncture
grumble	complain	gruvicé	grulus
sable	mink	saboured	sabet

APPENDIX B Supplementary Materials for Chapter 2

Table S2.1. Behavioural RTs analyses on all data versus data excluding items with pre-DP acoustic differences (identified in the gating post-test). Reported pairwise effects (planned) are one-tailed.

Contrast	All data			Data with exclusion		
	X^2	t	p	X^2	t	p
Lexicality	23.60		<.001	28.87		<.001
Lexicality-by-prime type	10.73		.005	8.52		.014
Word prime type	10.65		.005	8.57		.014
Word-word > word		3.33	<.001		3.00	.002
Word-word > pseudo-word		2.37	.009		2.30	.011
Pseudo prime type	1.62		.445	0.65		.720

Word-word, word-primed word; pseudo-word, pseudoword-primed word.

Table S2.2. Behavioural accuracy analyses on all data versus data excluding items with pre-DP acoustic differences. Reported pairwise effects are Bonferroni corrected.

Contrast	All data			Data with exclusion		
	X^2	z	p	X^2	z	p
Lexicality	7.31		.063	6.40		.094
Lexicality-by-prime type	6.08		.048	6.98		.031
Word prime type	13.95		<.001	14.97		<.001
Pseudo-word > word		3.14	.005		3.03	.007
Pseudo-word > word-word		3.07	.007		3.05	.007
Pseudo prime type	1.93		.381	3.16		.206

Word-word, word-primed word; pseudo-word, pseudoword-primed word.

Table S2.3. Pre-DP MEG analyses of unprimed > word-primed items and Post-DP MEG analyses of pseudoword > word on all data versus data excluding items with pre-DP acoustic differences. Reported effects are FWE corrected at cluster level at $p < 0.05$.

		All data			Data with exclusion		
Time		Cluster	Cluster	Latency	Cluster	Cluster	Latency
window	Modality	$P_{\text{FWE-corr}}$	size	(ms)	$P_{\text{FWE-corr}}$	size	(ms)
Pre-DP	Grad	.023	295	-28 to -4	.005	426	-25 to 0
Post-DP	Grad	<.001	39335	313 to 956	<.001	30811	320 to 775
	Mag	<.001	68517	359 to 990	<.001	69777	362 to 988
	Source	<.001	2315	400 to 900	<.001	2287	400 to 900

Grad, gradiometers; Mag, magnetometers.

Table S2.4. Pre-DP MEG ROI analyses on all data versus data excluding items with pre-DP acoustic differences across gradiometer sensor locations that showed post-DP pseudoword > word effect. Reported effects on unprimed > word-primed items (planned) are one-tailed.

Contrast	All data		Data with exclusion	
	<i>t</i>	<i>p</i>	<i>t</i>	<i>p</i>
Unprimed > word-primed	2.41	.013	2.57	.009
Unprimed > pseudo-primed	2.69	.014	3.14	.005

Unprimed, unprimed items; Word-primed, word-primed items; pseudo-primed, pseudoword-primed items.

Table S2.5. Post-DP MEG ROI analyses on all data versus data excluding items with pre-DP acoustic differences. Reported pairwise effects (planned) are one-tailed.

Contrast	Modality	All data			Data with exclusion		
		<i>F</i>	<i>t</i>	<i>p</i>	<i>F</i>	<i>t</i>	<i>p</i>
Lexicality-by-prime type	Grad	7.30		.002	6.12		.005
	Mag	5.80		.007	3.77		.035
	Source	0.99		.360	1.04		.354
Word prime type	Grad	8.01		.001	6.18		.005
	Mag	5.61		.009	4.46		.021
	Source	3.77		.038	3.64		.039
Word-word > word	Grad		2.22	.019		2.11	.023
	Mag		3.33	.002		2.79	.006
	Source		2.66	.007		2.51	.010
Word-word > pseudo-word	Grad		3.70	<.001		3.60	<.001
	Mag		2.64	.008		2.33	.015
	Source		1.26	.110		1.39	.089
Pseudo prime type	Grad	0.67		.514	0.57		.564
	Mag	0.80		.446	0.37		.681
	Source	1.12		.326	1.23		.300

Word-word, word-primed word; pseudo-word, pseudoword-primed word. Grad, gradiometers; Mag, magnetometers.

APPENDIX C Supplementary Materials for Chapter 3

Abbreviations for all the conditions:

W, unprimed word;

WW, word-primed word;

PW, pseudoword-primed word;

PPW, double pseudowords-primed word;

P, unprimed pseudoword;

PP, pseudoword-primed pseudoword;

WP, word-primed pseudoword;

WWP, double words-primed pseudoword

Table S3.1 Descriptive statistics for the data presented in Figure 3.3 (Experiment 2).

Conditions	Mean RTs (ms)	CI (ms)	Mean Error Rate	CI
W	732.499	22.113	0.095	0.023
WW	767.155	26.583	0.139	0.034
PW	743.953	26.577	0.120	0.028
PPW	750.835	17.390	0.120	0.025
P	852.037	28.138	0.065	0.022
PP	834.225	26.354	0.037	0.019
WP	850.007	24.358	0.041	0.020
WWP	853.951	22.814	0.044	0.024

Table S3.2 Descriptive statistics for the data presented in Figure 3.5 (Experiment 3).

Conditions	Mean RTs (ms)	CI (ms)	Mean Error Rate	CI
W	630.726	11.877	0.015	0.006
WW	638.244	14.953	0.012	0.008
PW	650.744	13.656	0.014	0.007
P	767.976	14.352	0.095	0.011
PP	770.320	15.514	0.072	0.013
WP	768.012	14.954	0.065	0.009

Table S3.3 Descriptive statistics for the data presented in Figure 3.8 (Experiment 4-5)

Experiment	Conditions	Mean RTs (ms)	CI (ms)	Mean Error Rate	CI
Exp4a (PD-LD)	W	707.200	11.631	0.127	0.015
	WW	718.529	10.943	0.141	0.019
	PW	713.051	11.986	0.144	0.016
	P	776.688	14.411	0.058	0.013
	PP	779.571	12.321	0.060	0.015
	WP	781.464	11.652	0.063	0.015
Exp4b (LD-LD)	W	727.937	10.480	0.149	0.016
	WW	743.559	15.474	0.161	0.017
	PW	734.252	12.271	0.179	0.018
	P	748.739	14.595	0.029	0.011
	PP	730.126	11.289	0.023	0.011
	WP	762.414	12.586	0.039	0.012
Exp5a (PD-PD)	W	660.657	11.098	0.022	0.007
	WW	641.152	9.798	0.018	0.006
	PW	656.087	12.665	0.021	0.007
	P	711.306	13.296	0.092	0.011
	PP	699.047	10.384	0.085	0.010
	WP	717.083	11.620	0.085	0.012
Exp5b (LD-PD)	W	681.047	9.899	0.030	0.008
	WW	684.827	12.292	0.026	0.007
	PW	672.660	10.144	0.026	0.008
	P	757.650	12.279	0.088	0.011
	PP	736.174	10.677	0.083	0.011
	WP	741.628	10.765	0.077	0.011

APPENDIX D Supplementary Materials for Chapter 4

Table S4.1 Descriptive statistics for the data presented in Figure 4.4 (Experiment 1).

Phase	Conditions	Mean Error Rate	CI	Mean Rating	CI
Phase1	W	0.157	0.028	1.617	0.065
	WW	0.133	0.025	1.616	0.055
	PW	0.169	0.030	1.650	0.067
	P	0.350	0.024	2.159	0.052
	PP	0.338	0.040	2.131	0.075
	WP	0.382	0.035	2.219	0.068
Phase2	W	0.174	0.034	1.673	0.075
	WW	0.131	0.024	1.631	0.075
	PW	0.213	0.039	1.740	0.074
	P	0.493	0.046	2.494	0.096
	PP	0.401	0.048	2.290	0.098
	WP	0.517	0.034	2.515	0.087
Both	W	0.146	0.027	1.607	0.063
Phases	WW	0.117	0.021	1.595	0.053
	PW	0.167	0.026	1.654	0.054
	P	0.414	0.027	2.311	0.059
	PP	0.361	0.040	2.194	0.081
	WP	0.439	0.024	2.348	0.050

Table S4.2 Descriptive statistics for the data presented in Figure 4.5 (Experiment 2).

Conditions	Mean Error Rate	CI	Mean Rating	CI
W	0.112	0.026	2.002	0.090
WW	0.112	0.022	1.991	0.085
PW	0.171	0.029	2.193	0.113
PPW	0.148	0.020	2.094	0.090
P	0.349	0.028	3.037	0.092
PP	0.401	0.028	3.169	0.094
WP	0.344	0.036	3.019	0.111
WWP	0.363	0.030	3.015	0.107

Table S4.3 Descriptive statistics for the data presented in Figure 4.6 (Experiment 3).

Conditions	Mean Error Rate	CI	Mean Rating	CI
W	0.250	0.020	2.592	0.068
WW	0.274	0.020	2.667	0.069
PW	0.278	0.024	2.716	0.077
P	0.410	0.021	3.195	0.073
PP	0.419	0.024	3.252	0.071
WP	0.404	0.021	3.210	0.065

Table S4.4 Descriptive statistics for the data presented in Figure 4.7 (Experiment 4-5)

Experiment	Conditions	Mean Error Rate	CI	Mean Rating	CI
Exp4a (PD-LD)	W	0.095	0.015	1.825	0.063
	WW	0.095	0.018	1.841	0.073
	PW	0.078	0.013	1.801	0.062
	P	0.311	0.020	2.903	0.065
	PP	0.323	0.020	2.919	0.082
	WP	0.333	0.020	2.923	0.067
Exp4b (LD-LD)	W	0.101	0.014	1.907	0.062
	WW	0.080	0.016	1.755	0.059
	PW	0.088	0.013	1.791	0.052
	P	0.357	0.021	3.002	0.059
	PP	0.357	0.024	3.007	0.082
	WP	0.339	0.023	2.928	0.081
Exp5a (PD-PD)	W	0.243	0.018	2.548	0.062
	WW	0.236	0.019	2.517	0.068
	PW	0.227	0.016	2.520	0.054
	P	0.435	0.020	3.324	0.058
	PP	0.439	0.020	3.310	0.066
	WP	0.407	0.023	3.241	0.067
Exp5b (LD-PD)	W	0.261	0.017	2.660	0.058
	WW	0.224	0.017	2.518	0.065
	PW	0.216	0.014	2.482	0.053
	P	0.424	0.021	3.285	0.069
	PP	0.427	0.022	3.293	0.064
	WP	0.425	0.023	3.300	0.069

BIBLIOGRAPHY

- Aitchison, L., & Lengyel, M. (2017). With or without you: predictive coding and Bayesian inference in the brain. *Current opinion in neurobiology*, 46, 219-227.
- Alba, J. W., & Hasher, L. (1983). Is memory schematic?. *Psychological Bulletin*, 93(2), 203.
- Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of memory and language*, 38(4), 419-439.
- Arndt, J., Lee, K., & Flora, D. B. (2008). Recognition without identification for words, pseudowords and nonwords. *Journal of memory and language*, 59(3), 346-360.
- Baayen, R. H., Piepenbrock, R., & van H, R. (1993). The {CELEX} lexical data base on {CD-ROM}.
- 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.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, 68(3), 255-278.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.
- Bayley, P. J., O'Reilly, R. C., Curran, T., & Squire, L. R. (2008). New semantic learning in patients with large medial temporal lobe lesions. *Hippocampus*, 18(6), 575-583.
- Bickel, D. R. (2015). Blending Bayesian and frequentist methods according to the precision of prior information with applications to hypothesis testing. *Statistical Methods & Applications*, 24(4), 523-546.

- Blank, H., Spangenberg, M., & Davis, M. H. (2018). Neural prediction errors distinguish perception and misperception of speech. *Journal of Neuroscience*, 38(27), 6076-6089.
- Blank, H., & Davis, M. H. (2016). Prediction errors but not sharpened signals simulate multivoxel fMRI patterns during speech perception. *PLoS biology*, 14(11), e1002577.
- Bozic, M., Tyler, L. K., Ives, D. T., Randall, B., & Marslen-Wilson, W. D. (2010). Bihemispheric foundations for human speech comprehension. *Proceedings of the National Academy of Sciences*, 107(40), 17439-17444.
- Breitenstein, C., Jansen, A., Deppe, M., Foerster, A. F., Sommer, J., Wolbers, T., & Knecht, S. (2005). Hippocampus activity differentiates good from poor learners of a novel lexicon. *Neuroimage*, 25(3), 958-968.
- Brodbeck, C., Hong, L. E., & Simon, J. Z. (2018). Rapid transformation from auditory to linguistic representations of continuous speech. *Current Biology*, 28(24), 3976-3983.
- Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. (2016). How many words do we know? Practical estimates of vocabulary size dependent on word definition, the degree of language input and the participant's age. *Frontiers in Psychology*, 7(JUL), 1–11.
- Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of statistical software*, 80(1), 1-28.
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., ... & Riddell, A. (2017). Stan: a probabilistic programming language. *Grantee Submission*, 76(1), 1-32.
- Christensen, R. H. B. (2015). ordinal—regression models for ordinal data. *R package version*, 28, 2015.

- Church, B. A., & Schacter, D. L. (1994). Perceptual specificity of auditory priming: implicit memory for voice intonation and fundamental frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(3), 521.
- Clark, D., & Wagner, A. D. (2003). Assembling and encoding word representations: fMRI subsequent memory effects implicate a role for phonological control. *Neuropsychologia*, 41(3), 304-317.
- Cleland, A. A., Gaskell, M. G., Quinlan, P. T., & Tamminen, J. (2006). Frequency effects in spoken and visual word recognition: Evidence from dual-task methodologies. *Journal of Experimental Psychology: Human Perception and Performance*, 32(1), 104.
- Cole, R. A. (1973). Listening for mispronunciations: A measure of what we hear during speech. *Perception & Psychophysics*, 13(1), 153-156.
- Corley, M., MacGregor, L. J., & Donaldson, D. I. (2007). It's the way that you, er, say it: Hesitations in speech affect language comprehension. *Cognition*, 105(3), 658-668.
- Cousineau, D. (2005). Confidence intervals in within-participant designs: A simpler solution to Loftus and Masson's method. *Tutorials in quantitative methods for psychology*, 1(1), 42-45.
- Craik, F. I., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General*, 125(2), 159.
- Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of verbal learning and verbal behavior*, 11(6), 671-684.
- Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of experimental Psychology: general*, 104(3), 268.

- Cycowicz, Y. M., Nessler, D., Horton, C., & Friedman, D. (2008). Retrieving object color: the influence of color congruity and test format. *Neuroreport*, 19(14), 1387.
- Dahan, D., Magnuson, J. S., & Tanenhaus, M. K. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive psychology*, 42(4), 317-367.
- Dahan, D., & Magnuson, J. S. (2006). Spoken word recognition. In *Handbook of psycholinguistics* (pp. 249-283). Academic Press.
- Davis, M.H. (2016). The Neurobiology of Lexical Access. In G. Hickok & S. L. Small (Eds.), *Neurobiology of language* (pp. 541-555). Academic Press.
- Davis, M. H., Di Betta, A. M., Macdonald, M. J., & Gaskell, M. G. (2009). Learning and consolidation of novel spoken words. *Journal of cognitive neuroscience*, 21(4), 803-820.
- Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: neural and behavioural evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1536), 3773-3800.
- Davis, M. H., Marslen-Wilson, W. D., & Gaskell, M. G. (2002). Leading up the lexical garden path: Segmentation and ambiguity in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 28(1), 218.
- Davis, M.H., Scharenborg, O. (2016) Speech perception by humans and machines. In Gaskell, G. & Mirkovic J. (Eds) *Speech Perception and Spoken Word Recognition*. Psychology Press.
- Davis, M. H. & Sohoglu E. (2020) Three Functions of Prediction Error for Bayesian Inference in Speech Perception. In: Gazzaniga, M., Mangun R., & Poeppel D. (Eds), *The Cognitive Neurosciences*, 6th Edition. MIT Press, Camb, MA, USA.

- de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, 47(1), 1-12. doi:10.3758/s13428-014-0458-y.
- Desroches, A. S., Newman, R. L., & Joanisse, M. F. (2009). Investigating the time course of spoken word recognition: Electrophysiological evidence for the influences of phonological similarity. *Journal of Cognitive Neuroscience*, 21(10), 1893-1906.
- Donhauser, P. W., & Baillet, S. (2020). Two distinct neural timescales for predictive speech processing. *Neuron*, 105(2), 385-393.
- Dufour, S., & Peereman, R. (2003). Inhibitory priming effects in auditory word recognition: When the target's competitors conflict with the prime word. *Cognition*, 88(3), B33-B44.
- Dufour, S., & Peereman, R. (2003). Lexical competition in phonological priming: Assessing the role of phonological match and mismatch lengths between primes and targets. *Memory & Cognition*, 31(8), 1271-1283.
- Dufour, S., Brunellière, A., & Frauenfelder, U. H. (2013). Tracking the time course of word-frequency effects in auditory word recognition with event-related potentials. *Cognitive science*, 37(3), 489-507.
- Dufour, S., & Nguyen, N. (2017). Does talker-specific information influence lexical competition? Evidence from phonological priming. *Cognitive science*, 41(8), 2221-2233.
- Dumay, N., & Gaskell, M. G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science*, 18(1), 35-39.
- Dumay, N., & Gaskell, M. G. (2012). Overnight lexical consolidation revealed by speech segmentation. *Cognition*, 123(1), 119-132.

- Eberhard, K. M. (1994). Phonological inhibition in auditory word recognition. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention memory and language*. San Diego, CA: Academic Press.
- Fernandes, T., Kolinsky, R., & Ventura, P. (2009). The metamorphosis of the statistical segmentation output: Lexicalization during artificial language learning. *Cognition*, 112(3), 349-366.
- Ferreira, F., & Patson, N. D. (2007). The 'good enough' approach to language comprehension. *Language and Linguistics Compass*, 1(1-2), 71-83.
- Forbach, G. B., Stanners, R. F., & Hochhaus, L. (1974). Repetition and practice effects in a lexical decision task. *Memory & Cognition*, 2(2), 337-339.
- Frauenfelder, U. H., & Tyler, L. K. (1987). The process of spoken word recognition: An introduction. *Cognition*, 25(1-2), 1-20.
- Gagnepain, P., Lebreton, K., Desgranges, B., & Eustache, F. (2008). Perceptual priming enhances the creation of new episodic memories. *Consciousness and Cognition*, 17(1), 276-287.
- Gagnepain, P., Henson, R. N., & Davis, M. H. (2012). Temporal predictive codes for spoken words in auditory cortex. *Current Biology*, 22(7), 615-621.
- Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of experimental psychology: Human perception and performance*, 6(1), 110.
- Gaskell, M. G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition*, 89(2), 105-132.
- Gaskell, M. G., & Marslen-Wilson, W. D. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and cognitive Processes*, 12(5-6), 613-656.

- Gaskell, M. G., & Marslen-Wilson, W. D. (1999). Ambiguity, competition, and blending in spoken word recognition. *Cognitive Science*, 23(4), 439-462.
- Gaskell, M. G., & Marslen-Wilson, W. D. (2002). Representation and competition in the perception of spoken words. *Cognitive psychology*, 45(2), 220-266.
- Goldinger, S. D. (1996). Words and voices: episodic traces in spoken word identification and recognition memory. *Journal of experimental psychology: Learning, memory, and cognition*, 22(5), 1166.
- Goldinger, S. D., Luce, P. A., Pisoni, D. B., & Marcario, J. K. (1992). Form-based priming in spoken word recognition: The roles of competition and bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(6), 1211.
- Green, P., & MacLeod, C. J. (2016). SIMR: an R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, 7(4), 493-498.
- Greve, A., Cooper, E., Kaula, A., Anderson, M. C., & Henson, R. (2017). Does prediction error drive one-shot declarative learning?. *Journal of memory and language*, 94, 149-165.
- Greve, A., Cooper, E., Tibon, R., & Henson, R. N. (2019). Knowledge is power: Prior knowledge aids memory for both congruent and incongruent events, but in different ways. *Journal of Experimental Psychology: General*, 148(2), 325.
- Gronau, Q. F., Sarafoglou, A., Matzke, D., Ly, A., Boehm, U., Marsman, M., ... & Steingroever, H. (2017). A tutorial on bridge sampling. *Journal of mathematical psychology*, 81, 80-97.
- Gwilliams, L.E., & Davis, M.H. (2021) Extracting language content from speech sounds: the information theoretic approach. In L.L. Holt, J.E. Peelle, A.B. Coffin, A.N. Popper, R.R. Fay (Eds.), *The Auditory cognitive neuroscience of speech perception: Springer handbook of auditory research*. New York: Springer.

- Haeuser, K. I., & Kray, J. (2021). Effects of prediction error on episodic memory retrieval: evidence from sentence reading and word recognition. *Language, Cognition and Neuroscience*, 1-17.
- Henson, R. N., & Gagnepain, P. (2010). Predictive, interactive multiple memory systems. *Hippocampus*, 20(11), 1315-1326.
- Henson, R.N., Wakeman, D.G., Litvak, V. & Friston, K.J. (2011). A Parametric Empirical Bayesian framework for the EEG/MEG inverse problem: generative models for multiparticipant and multimodal integration. *Frontiers in Human Neuroscience*, 5, 76, 1-16.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature reviews neuroscience*, 8(5), 393-402.
- Höltje, G., Lubahn, B., & Mecklinger, A. (2019). The congruent, the incongruent, and the unexpected: Event-related potentials unveil the processes involved in schematic encoding. *Neuropsychologia*, 131, 285-293.
- Howes, D. H., & Solomon, R. L. (1951). Visual duration threshold as a function of word-probability. *Journal of experimental psychology*, 41(6), 401.
- Huetting, F. (2015). Four central questions about prediction in language processing. *Brain research*, 1626, 118-135.
- James, E., Gaskell, M. G., Weighall, A., & Henderson, L. (2017). Consolidation of vocabulary during sleep: The rich get richer?. *Neuroscience & Biobehavioral Reviews*, 77, 1-13.
- Jurica, P. J., & Shimamura, A. P. (1999). Monitoring item and source information: Evidence for a negative generation effect in source memory. *Memory & cognition*, 27(4), 648-656.

- Kafkas, A., & Montaldi, D. (2015). Striatal and midbrain connectivity with the hippocampus selectively boosts memory for contextual novelty. *Hippocampus*, 25(11), 1262-1273.
- Kapnoula, E. C., Packard, S., Gupta, P., & McMurray, B. (2015). Immediate lexical integration of novel word forms. *Cognition*, 134, 85-99.
- Kapur, N., & Brooks, D. J. (1999). Temporally-specific retrograde amnesia in two cases of discrete bilateral hippocampal pathology. *Hippocampus*, 9(3), 247-254.
- Kapur, S., Craik, F. I., Tulving, E., Wilson, A. A., Houle, S., & Brown, G. M. (1994). Neuroanatomical correlates of encoding in episodic memory: Levels of processing effect. *Proceedings of the National Academy of Sciences*, 91(6), 2008-2011.
- Kawahara, H. (2006). STRAIGHT, exploitation of the other aspect of VOCODER: Perceptually isomorphic decomposition of speech sounds. *Acoustical science and technology*, 27(6), 349-353.
- Kilner, J. M., & Friston, K. J. (2010). Topological inference for EEG and MEG. *The Annals of Applied Statistics*, 1272-1290.
- Kim, H. (2011). Neural activity that predicts subsequent memory and forgetting: a meta-analysis of 74 fMRI studies. *Neuroimage*, 54(3), 2446-2461.
- Kisler, T. and Reichel U. D. and Schiel, F. (2017): Multilingual processing of speech via web services, *Computer Speech & Language*, Volume 45, September 2017, pages 326–347.
- Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3?.
- Kocagoncu, E., Clarke, A., Devereux, B. J., & Tyler, L. K. (2017). Decoding the cortical dynamics of sound-meaning mapping. *Journal of Neuroscience*, 37(5), 1312-1319.
- Kumaran, D., Hassabis, D., & McClelland, J. L. (2016). What learning systems do intelligent agents need? Complementary learning systems theory updated. *Trends in cognitive sciences*, 20(7), 512-534.

- Landi, N., Malins, J. G., Frost, S. J., Magnuson, J. S., Molfese, P., Ryherd, K., ... & Pugh, K. R. (2018). Neural representations for newly learned words are modulated by overnight consolidation, reading skill, and age. *Neuropsychologia*, *111*, 133-144.
- Lange, K., Kühn, S., & Filevich, E. (2015). " Just Another Tool for Online Studies"(JATOS): An Easy Solution for Setup and Management of Web Servers Supporting Online Studies. *PloS one*, *10*(6), e0130834.
- Li, W., Qin, W., Liu, H., Fan, L., Wang, J., Jiang, T., & Yu, C. (2013). Subregions of the human superior frontal gyrus and their connections. *Neuroimage*, *78*, 46-58.
- Litvak, V., Mattout, J., Kiebel, S., Phillips, C., Henson, R., Kilner, J., ... & Penny, W. (2011). EEG and MEG data analysis in SPM8. *Computational intelligence and neuroscience*, *2011*.
- Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and hearing*, *19*(1), 1.
- Luthra, S., Li, M. Y., You, H., Brodbeck, C., & Magnuson, J. S. (2021). Does signal reduction imply predictive coding in models of spoken word recognition?. *Psychonomic Bulletin & Review*, 1-9.
- Magnuson, J. S., Mirman, D., & Harris, H. D. (2012). Computational models of spoken word recognition. *The Cambridge handbook of psycholinguistics*, 76-103.
- Mandera, P., Keuleers, E., & Brysbaert, M. (2017). Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language*, *92*, 57-78.
- Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive psychology*, *10*(1), 29-63.

- Marslen-Wilson WD. 1984. Function and process in spoken word recognition. In: Bouma H, Bouwhuis DG, editors. Attention and performance X: control of language processes. Hillsdale (NJ): Erlbaum. p. 125–150.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25(1-2), 71-102.
- Marslen-Wilson, W. (1990). Activation, competition, and frequency in lexical access. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 148–172). The MIT Press.
- Marslen-Wilson, W. (1993). Issues of process and representation in lexical access. In G. T. M. Altmann & R. Shillcock (Eds.), *Cognitive models of speech processing: The Second Sperlonga Meeting* (pp. 187–210). Lawrence Erlbaum Associates Publishers.
- Marslen-Wilson, W., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: words, phonemes, and features. *Psychological review*, 101(4), 653.
- Marsolek, C. J. (2008). What antipriming reveals about priming. *Trends in Cognitive Sciences*, 12(5), 176-181.
- Martins, S., Guillery-Girard, B., Jambaqué, I., Dulac, O., & Eustache, F. (2006). How children suffering severe amnesic syndrome acquire new concepts?. *Neuropsychologia*, 44(14), 2792-2805.
- Mattout, J., Henson, R. N., and Friston, K. J. (2007). Canonical source reconstruction for MEG. *Comp. Int. Neurosci.* 2007, 67613.
- Mattys, S. L., & Clark, J. H. (2002). Lexical activity in speech processing: Evidence from pause detection. *Journal of Memory and Language*, 47(3), 343-359.
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive psychology*, 18(1), 1-86.

- McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102, 419–457.
- Mckone, E., & Trynes, K. (1999). Acquisition of novel traces in short-term implicit memory: Priming for nonwords and new associations. *Memory & cognition*, 27(4), 619-632.
- McMurray, B., Kapnoula, E. C., & Gaskell, M. G. (2016). Learning and integration of new word-forms: Consolidation, pruning, and the emergence of automaticity. In *Speech perception and spoken word recognition* (pp. 126-152). Psychology Press.
- Mestres-Missé, A., Camara, E., Rodriguez-Fornells, A., Rotte, M., & Münte, T. F. (2008). Functional neuroanatomy of meaning acquisition from context. *Journal of Cognitive Neuroscience*, 20(12), 2153-2166.
- Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they?. *Journal of Experimental Psychology: General*, 118(1), 43.
- Monsell, S., & Hirsh, K. W. (1998). Competitor priming in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(6), 1495.
- Newman, S. D., & Twieg, D. (2001). Differences in auditory processing of words and pseudowords: An fMRI study. *Human brain mapping*, 14(1), 39-47.
- Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, 52(3), 189-234.
- Norris, D., Cutler, A., McQueen, J. M., & Butterfield, S. (2006). Phonological and conceptual activation in speech comprehension. *Cognitive Psychology*, 53(2), 146-193.

- Norris, D., & Kinoshita, S. (2008). Perception as evidence accumulation and Bayesian inference: Insights from masked priming. *Journal of Experimental Psychology: General*, 137(3), 434.
- Norris, D., & McQueen, J. M. (2008). Shortlist B: a Bayesian model of continuous speech recognition. *Psychological review*, 115(2), 357.
- Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, 23(3), 299-325.
- Norris, D., McQueen, J. M., & Cutler, A. (2002). Bias effects in facilitatory phonological priming. *Memory & Cognition*, 30(3), 399-411.
- Okada, K., & Hickok, G. (2006). Identification of lexical–phonological networks in the superior temporal sulcus using functional magnetic resonance imaging. *Neuroreport*, 17(12), 1293-1296.
- Orfanidou, E., Marslen-Wilson, W. D., & Davis, M. H. (2006). Neural response suppression predicts repetition priming of spoken words and pseudowords. *Journal of cognitive neuroscience*, 18(8), 1237-1252.
- Orfanidou, E., Davis, M. H., Ford, M. A., & Marslen-Wilson, W. D. (2011). Perceptual and response components in repetition priming of spoken words and pseudowords. *The Quarterly Journal of Experimental Psychology*, 64(1), 96-121.
- O'Rourke, T. B., & Holcomb, P. J. (2002). Electrophysiological evidence for the efficiency of spoken word processing. *Biological psychology*, 60(2-3), 121-150.
- Otten, L. J., Sveen, J., & Quayle, A. H. (2007). Distinct patterns of neural activity during memory formation of nonwords versus words. *Journal of Cognitive Neuroscience*, 19(11), 1776-1789.
- Paller, K. A., & Wagner, A. D. (2002). Observing the transformation of experience into memory. *Trends in cognitive sciences*, 6(2), 93-102.

- Plaut, D. C. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language and cognitive processes*, 12(5-6), 765-806.
- Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage*, 10(1), 15-35.
- Prabhakaran, R., Blumstein, S. E., Myers, E. B., Hutchison, E., & Britton, B. (2006). An event-related fMRI investigation of phonological–lexical competition. *Neuropsychologia*, 44(12), 2209-2221.
- Quent, J. A., Henson, R. N., & Greve, A. (2021). A predictive account of how novelty influences declarative memory. *Neurobiology of Learning and Memory*, 179, 107382.
- Radeau, M., Morais, J., & Segui, J. (1995). Phonological priming between monosyllabic spoken words. *Journal of Experimental Psychology: Human Perception and Performance*, 21(6), 1297.
- Reichardt, R., Polner, B., & Simor, P. (2020). Novelty manipulations, memory performance, and predictive coding: The role of unexpectedness. *Frontiers in human neuroscience*, 14, 152.
- Rock, I. (1957). The role of repetition in associative learning. *The American journal of psychology*, 70(2), 186-193.
- Rodd, J. M., Cai, Z. G., Betts, H. N., Hanby, B., Hutchinson, C., & Adler, A. (2016). The impact of recent and long-term experience on access to word meanings: Evidence from large-scale internet-based experiments. *Journal of Memory and Language*, 87, 16-37.

- Schacter, D. L., Church, B., & Treadwell, J. (1994). Implicit memory in amnesic patients: Evidence for spared auditory priming. *Psychological Science*, 5(1), 20-25.
- Schiel, F. (1999). Automatic Phonetic Transcription of Non-Prompted Speech. In Proc. of the ICPhS (pp. 607-610).
- Schönbrodt, F. D. & Stefan, A. M. (2018). BFDA: An R package for Bayes factor design analysis (version 0.3). Retrieved from <https://github.com/nicebread/BFDA>
- Schönbrodt, F. D. & Wagenmakers, E.-J. (2018). Bayes Factor Design Analysis: Planning for compelling evidence. *Psychonomic Bulletin & Review*, 128-142.
- Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological review*, 96(4), 523.
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, 27(3), 379-423.
- Shtyrov, Y., Smith, M. L., Horner, A. J., Henson, R., Nathan, P. J., Bullmore, E. T., & Pulvermüller, F. (2012). Attention to language: novel MEG paradigm for registering involuntary language processing in the brain. *Neuropsychologia*, 50(11), 2605-2616.
- Slowiaczek, L. M., & Hamburger, M. (1992). Prelexical facilitation and lexical interference in auditory word recognition. *Journal of experimental psychology: Learning, memory, and cognition*, 18(6), 1239.
- Slowiaczek, L. M., & Pisoni, D. B. (1986). Effects of phonological similarity on priming in auditory lexical decision. *Memory & cognition*, 14(3), 230-237.
- Sohoglu, E., & Davis, M. H. (2016). Perceptual learning of degraded speech by minimizing prediction error. *Proceedings of the National Academy of Sciences*, 113(12), E1747-E1756.

- Sohoglu, E., Peelle, J. E., Carlyon, R. P., & Davis, M. H. (2012). Predictive top-down integration of prior knowledge during speech perception. *Journal of Neuroscience*, 32(25), 8443-8453.
- Spratling, M. W. (2008). Reconciling predictive coding and biased competition models of cortical function. *Frontiers in computational neuroscience*, 2, 4.
- Squire, L. R. (2009). The legacy of patient HM for neuroscience. *Neuron*, 61(1), 6-9.
- Staresina, B. P., Gray, J. C., & Davachi, L. (2009). Event congruency enhances episodic memory encoding through semantic elaboration and relational binding. *Cerebral Cortex*, 19(5), 1198-1207.
- Stefan, A. M., Gronau, Q. F., Schönbrodt, F. D., & Wagenmakers, E. (2018). A tutorial on Bayes Factor Design Analysis using an informed prior. *Behavior Research Methods*, 1042-1058.
- Sweegers, C. C. G., Coleman, G. A., van Poppel, E. A. M., Cox, R., & Talamini, L. M. (2015). Mental schemas hamper memory storage of goal-irrelevant information. *Frontiers in human neuroscience*, 9, 629.
- Takashima, A., Bakker, I., Van Hell, J. G., Janzen, G., & McQueen, J. M. (2014). Richness of information about novel words influences how episodic and semantic memory networks interact during lexicalization. *NeuroImage*, 84, 265-278.
- Takashima, A., Bakker, I., Van Hell, J. G., Janzen, G., & McQueen, J. M. (2017). Interaction between episodic and semantic memory networks in the acquisition and consolidation of novel spoken words. *Brain and language*, 167, 44-60.
- Tamminen, J., Payne, J. D., Stickgold, R., Wamsley, E. J., & Gaskell, M. G. (2010). Sleep spindle activity is associated with the integration of new memories and existing knowledge. *Journal of Neuroscience*, 30(43), 14356-14360.

- Tauroza, S., & Allison, D. (1990). Speech rates in british english. *Applied linguistics*, 11(1), 90-105.
- Taylor, J. S. H., Rastle, K., & Davis, M. H. (2013). Can cognitive models explain brain activation during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. *Psychological bulletin*, 139(4), 766.
- Taylor, J. S. H., Rastle, K., & Davis, M. H. (2014). Interpreting response time effects in functional imaging studies. *Neuroimage*, 99, 419-433.
- Tulving, E., & Kroll, N. (1995). Novelty assessment in the brain and long-term memory encoding. *Psychonomic Bulletin & Review*, 2(3), 387-390.
- van Casteren, M., & Davis, M. H. (2006). Mix, a program for pseudorandomization. *Behavior research methods*, 38(4), 584-589.
- van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, 67(6), 1176-1190.
- Van Kesteren, M. T., Ruiter, D. J., Fernández, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in neurosciences*, 35(4), 211-219.
- Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological science*, 9(4), 325-329.
- Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of memory and language*, 40(3), 374-408.
- Vitevitch, M. S., Luce, P. A., Pisoni, D. B., & Auer, E. T. (1999). Phonotactics, neighborhood activation, and lexical access for spoken words. *Brain and language*, 68(1-2), 306-311.

- Vitevitch, M. S., & Luce, P. A. (2005). Increases in phonotactic probability facilitate spoken nonword repetition. *Journal of memory and language*, 52(2), 193-204.
- Wagenmakers, E. J., Lodewyckx, T., Kuriyal, H., & Grasman, R. (2010). Bayesian hypothesis testing for psychologists: A tutorial on the Savage–Dickey method. *Cognitive psychology*, 60(3), 158-189.
- Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., ... & Buckner, R. L. (1998). Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. *Science*, 281(5380), 1188-1191.
- Wang, Y. C., Sohoglu, E., Gilbert, R. A., Henson, R. N., & Davis, M. H. (2021). Predictive Neural Computations Support Spoken Word Recognition: Evidence from MEG and Competitor Priming. *Journal of Neuroscience*.
- Woods, D. L., Wyma, J. M., Yund, E. W., Herron, T. J., & Reed, B. (2015). Factors influencing the latency of simple reaction time. *Frontiers in human neuroscience*, 9, 131.
- Yi, H. G., Leonard, M. K., & Chang, E. F. (2019). The encoding of speech sounds in the superior temporal gyrus. *Neuron*, 102(6), 1096-1110.
- Ylinen, S., Bosseler, A., Junttila, K., & Huotilainen, M. (2017). Predictive coding accelerates word recognition and learning in the early stages of language development. *Developmental science*, 20(6), e12472.
- Yonelinas, A. P., Ranganath, C., Ekstrom, A. D., & Wiltgen, B. J. (2019). A contextual binding theory of episodic memory: systems consolidation reconsidered. *Nature Reviews Neuroscience*, 20(6), 364-375.

Zhuang, J., Randall, B., Stamatakis, E. A., Marslen-Wilson, W. D., & Tyler, L. K. (2011).

The interaction of lexical semantics and cohort competition in spoken word recognition: an fMRI study. *Journal of Cognitive Neuroscience*, 23(12), 3778-3790.

Zhuang, J., Tyler, L. K., Randall, B., Stamatakis, E. A., & Marslen-Wilson, W. D. (2014).

Optimally efficient neural systems for processing spoken language. *Cerebral Cortex*, 24(4), 908-918.