Homophone Representation in Monolingual and Bilingual Impaired and Unimpaired Speakers

Polly Barr

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

ARC Centre of Excellence in Cognition and its Disorders (CCD)

Department of Cognitive Science

Macquarie University, Sydney, Australia

November 2016

Declaration

The work in this thesis is my own original work. It has not been submitted for a higher degree in any other university or institution. All of the work reported in this thesis was undertaken during the time I was enrolled as a PhD student at Macquarie University under the supervision of Prof. Lyndsey Nickels, Dr. Britta Biedermann and Dr. Marie-Josèphe Tainturier. Ethics approval for the studies reported in this thesis was obtained from Macquarie University's Human Research Ethics Committee, Reference No. 5201200905 HE27707RO5342 and 10/WNo01/67.

Signed:

Polly Barr, 18th November 2016

Table of contents

General Abstract

Previous research has suggested three possible architectures for homophone production. Homophones may be stored either with shared representations, producing a processing advantage relative to non-homophones, or with independent representations and feedforward activation, producing no benefit for homophones, or with independent representations and interactive activation flow, where homophones may or may not have a production advantage depending on the relative balance between the influence of processing stages in retrieval. The purpose of this thesis was to attempt to reconcile the previously diverse findings in the literature and investigate which account was the most plausible theory regarding homophone production in the phonological and orthographic output lexicons.

The first experimental chapter investigated whether previous conflicting evidence was due to differences in participants (monolingual vs bilingual), tasks (picture naming vs translation) or spelling (heterographs vs homographs). It did this by examining picture naming and translation of homophones and controls matched to their individual or summed frequency with monolingual and bilingual speakers.

The second experimental chapter aimed to replicate phonological homophone treatment generalisation in the treatment of spelling with a participant with acquired dysgraphia. This treatment methodology examined whether improved lexical retrieval following treatment of one homophone resulted in generalisation to written naming of its untreated partner. Another aim of this experiment was to investigate if generalisation (if any) was due to homophones (homographs) sharing an orthographic word form or due to feedback from treated graphemes to independent

representations. Therefore generalisation to items with varying degrees of orthographic overlap was also investigated (heterographs and direct neighbours).

The final experimental chapter examined the effect of homograph priming compared to repetition, heterograph and direct neighbour priming in written picture naming.

The results across these experiments provided no support for shared representations. Instead, the pattern of results favoured an account with independent representations and interactive activation including feedback from segments to the lexicon, however further research and computational modelling is needed to fully support this hypothesis.

Acknowledgements

I would like to express my deepest gratitude to my supervisors as they have gone above and beyond the call of duty in mentoring me for this thesis. Everything produced in this thesis is a testament to their patience, brilliance and excellent supervision. I could not have been luckier to have been supervised by such inspirational (both academically and personally) supervisors and just hope that one day I am half the academic and supervisor you all are (but by gum you've set the bar high!). Obviously I could not have done this without them and I am extremely grateful for everything they have taught me- I would not be the academic I am today without them.

I would also like to thank my team of friends and colleagues on both sides of the world who helped me get through not only the last couple of weeks but through the past 3.5 years who have kept me sane. Again I'm very grateful and would not be who I am without you. In particular my little Australian family-thank-you for making this feel like home when I felt lost. For the same reason I would like to thank my wonderful family who have managed to give amazing emotional support from afar (in particular my mother father and brothers). I am incredibly lucky to have such a wonderful family.

I would like to thank all my participants (especially the 'pilot' participants i.e. unwilling friends and family, I roped in with minimal reward!) including my hard working patient- who is possibly the most motivated person I have ever met! Finally, I am great-ful for being able to spend the past three and a half years making homophone puns.

Chapter 1: Introduction

Sewing the Storey

This thesis focuses on language production, both spoken and written, examining the nature of representation and processing. This chapter provides a general background to the thesis outlining relevant language production models, and in particular discussing debates related to the steps required to access phonology from semantics.

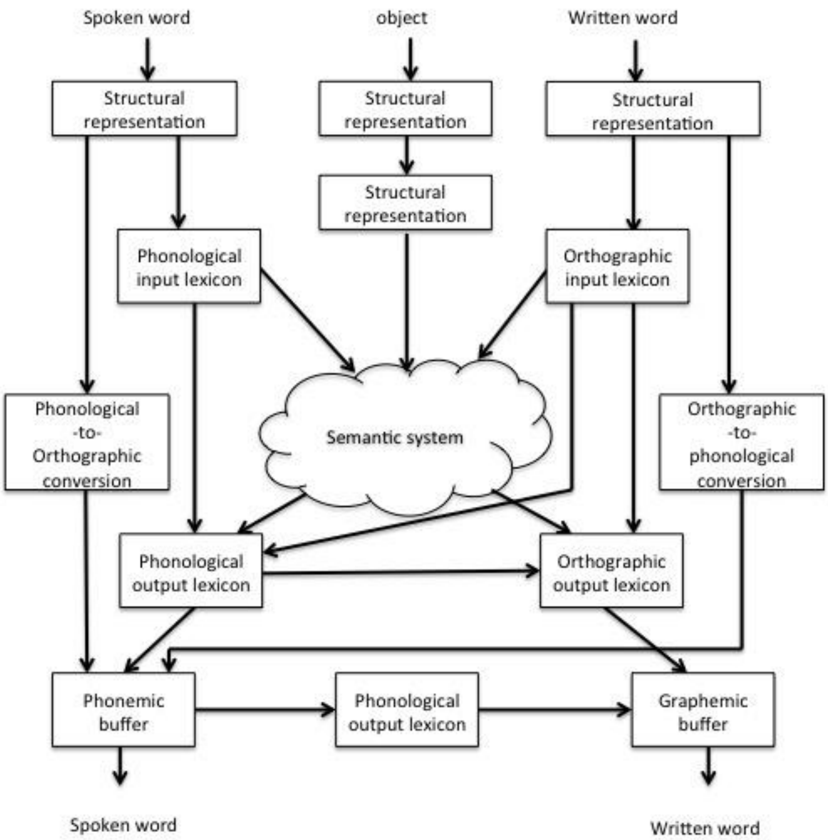


Figure 1. A sketch of the processes for processing of a single word (adapted from Patterson, 1986).

Even processing a single word is a potentially complex process. Figure 1 portrays the various routes through the lexicon that may be used in order to understand or produce a word. The choice of path depends on the nature of the task, various properties of the word, previous experience with that item, as well as the hurdles introduced if a speaker has a developmental or acquired language impairment. Nevertheless, there is general consensus regarding the broad order of and nature of processing steps as illustrated in Figure 1.

If we think about producing a single word like 'cat', this can be either spoken aloud or written down. In Figure 1, this is reflected in there being two output routes. Both start with activation of the semantic system (also known as lexical semantics, or lexical concepts) - a modality-neutral store of word meanings. For production this is activated by a thought, idea, or object in the environment. The semantic system links to separate phonological (for spoken production) and orthographic (for written production) lexicons (e.g., Rapp & Caramazza, 1997; Wilshire, 2008; Wilshire & McCarthy, 2002)[1]. These separate (spoken and written) output lexicons will be a focus of this thesis. This chapter will first discuss the process of spoken word production followed by written word production.

Throughout this thesis there is a focus on homophone processing. Homophones are words that share the same phonological form despite having different meanings (e.g. *bat*-sports equipment, and *bat*-flying mammal). Homophones may share both phonology and spelling (homographs; e.g. *band*-group and *band*-loop) or share phonology but not spelling (heterographs; e.g., *key* and *quay*). These unique characteristics mean that homophones are an ideal resource to investigate access of word forms from semantics, but also that they pose challenges regarding their

[1] In Figure 1, separate phonological (and orthographic) input and output lexicons are depicted. However, other authors propose a single, shared lexicon used for both input and output (e.g. Jones & Rawson, 2016). However, as, in this thesis, the focus is on production of a word from meaning (via picture naming), whether the input and output lexicons are separate or not does not impact on the discussion and hence this thesis takes a neutral stance.

representation. This chapter will therefore also briefly discuss the different methods used to investigate homophone representation, however these are discussed in further detail in each experimental chapter. Finally, this introductory chapter provides an overview of the structure of the rest of the thesis.

Spoken Production

Virtually all models of spoken word production, whether this production begins with an idea, object, picture or definition, involve sequentially accessing semantics, and then phonology (i.e., first activating the semantic system and then the phonological output lexicon in Figure 1). Following this mapping of meaning onto a phonological form the individual phonemes are spelled out, leading ultimately to articulation. There is also recognition that information about syntactic properties of the word (e.g. word class, grammatical-gender, grammatical number, tense, etc.) must be stored. However, theories differ in how they represent lexical-syntax. They also differ in the nature of activation flow between processing levels.

There are three main models of spoken production, and these are outlined below: Levelt's Two-Stage model of lexical access (Levelt, 1999; Levelt, Roelofs, & Meyer, 1999; Roelofs, 1997), Caramazza's Independent Network model (Caramazza, 1997; Caramazza, Costa, Miozzo, & Bi, 2001) and Dell's Interactive Activation model (Dell, 1986, 1990; Dell, Chang, & Griffin, 1999; Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Schwartz, Dell, Martin, Gahl, & Sobel, 2006). Despite some similarities, these three models differ in the steps and methods of activation to retrieve phonological word forms from lexical-semantics. Moreover, these theories differ in the mechanisms proposed to account for frequency effects. Frequency effects refer to the fact that words that are more common in the speaker's experience are accessed easier. For example, Dell (1990) found, across various experiments, that participants made fewer errors when completing sentences with high frequency words compared to low frequency ones. Frequency effects are of particular relevance to this thesis because, as discussed below, they have been critical to the debate on the representation of homophones.

The Two-Stage model of lexical access

Levelt et al.'s Two-Stage (1999) has been one of the most influential theories in the literature. One of the distinctive features of this model is that lexicalisation occurs in two separate stages after a lexical-semantic entry has been selected. Each concept has a link to a representation that also points to information regarding that concept's lexical-syntax - the lemma. For example, for the lexical concept of KEY the corresponding lemma would specify that it is a count noun, its grammatical gender (language dependent, e.g. masculine in German, feminine in Spanish) and its number (whether it is singular or plural). Each lemma has a unique pointer to a specific phonological code. This unique link means that only if the corresponding lemma node is activated will the phonological word form node be activated. Activating the lemma is the first stage of lexicalisation. Following selection of the most highly activated lemma, the phonological node corresponding to the chosen lemma is activated at the word form level (the second stage of lexicalisation, see Figure 2). The lemma and phonological word form are accessed strictly sequentially. Finally, after the phonological word form has been selected, activation spreads to individual phonemes to be prepared for articulation.

It needs to be noted that this is a strictly feedforward model, hence there is no interaction between the lemma and word form or word form and phoneme levels. In addition, while there is parallel activation of lemmas, once a lemma has been selected only activation from this lemma (and unique pointer) feeds forward to one single word node, resulting in only one word node being active at any one time.

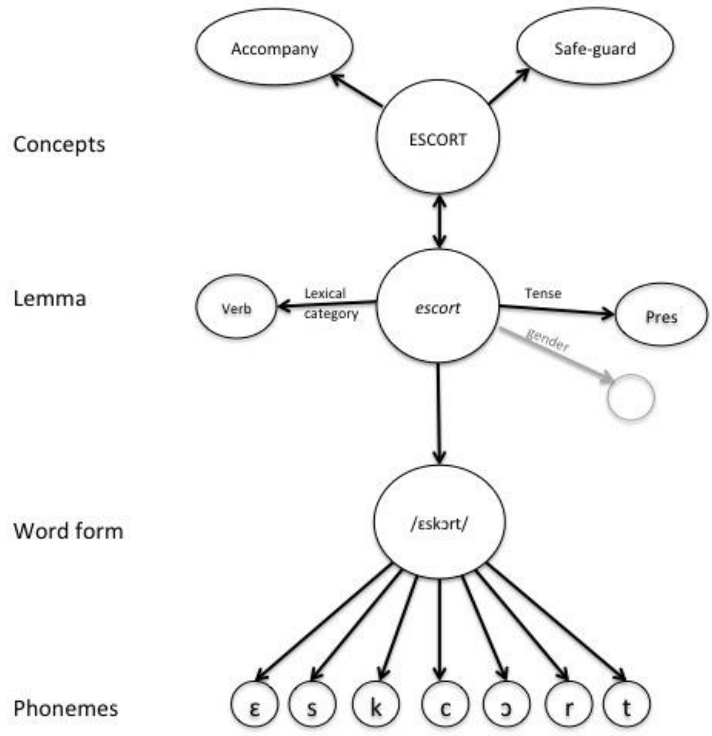


Figure 2. Adaptation of the Two-Stage model of lexical access, taken from Levelt et al. (1999).

Levelt et al. (1999) argue that converging evidence from latency and speech error data supports sequential lemma and word form levels. One strand of evidence stems from tip-of-the-tongue (TOT) experiments. Vigliocco, Antonini, and Garrett (1997) found that Italian speakers were able to produce the gender of a word (requiring lemma access) but not its phonological form when in a TOT state. This was further supported by evidence that an Italian participant with aphasia could nearly always produce the target word's grammatical gender when unable to produce the word form but never the reverse (Badecker, Miozzo, & Zanuttini's 1995). These studies alongside speech errors in non-impaired participants and reaction time studies (e.g., Jescheniak & Levelt, 1994) were used to support the syntactic mediation hypothesis: a lemma is stored separately to its word form and is retrieved prior to access to phonology (see for example, Roelofs, Meyer, & Levelt, 1998).

Frequency effects within Levelt et al.'s (1999) framework are argued to arise at the word form level on the basis of data from a series of experiments examining word production by Jescheniak and Levelt (1994).

In sum, Levelt et al.'s (1999) Two-Stage model is a feedforward, serial model for spoken word production. It proposes two stages to word form retrieval. Firstly, a lemma is activated and selected, followed by activation of a (frequency-sensitive) phonological word form.

The Independent Network model

Like other lexical access models, the Independent Network model (Caramazza et al., 2001) represents lexical-semantic information separately to lexical-syntax and word form information. The major difference between the Independent Network model and the Two-Stage model is that in the Independent Network model syntactic information is stored at the same level as phonological information as shown in Figure 3 (Caramazza et al., 2001, Caramazza & Miozzo, 1997; Miozzo & Caramazza, 1997). Miozzo and Caramazza (1997) postulate that a more parsimonious assumption for lexicalisation is to assume one lexical level with two independent lexical-syntactic and phonological networks linked to the semantic system. The main route from semantics to articulation occurs through the phonological network, therefore access to the phonological word form does not depend on previous access to lexical-syntax.

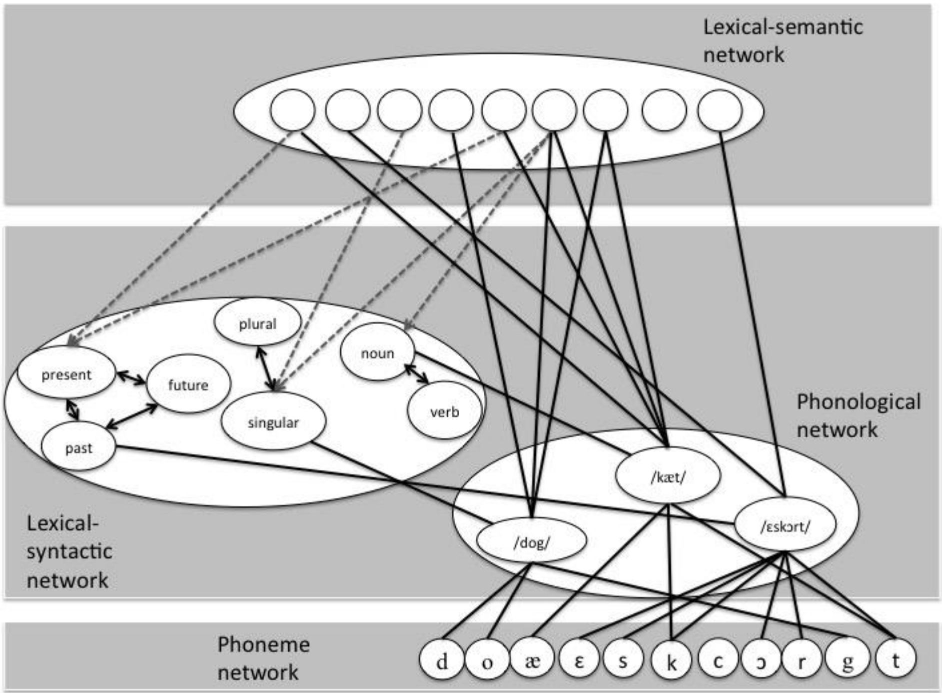


Figure 3. Adaptation of the Independent Network model (Caramazza, 1997).

Miozzo and Caramazza (1997b) argue that Levelt et al.'s (1999) syntactic mediation account predicted that participants who experience a TOT state should recall significantly more information regarding a word's grammatical properties if they are able to recall phonological properties, yet, this is not supported by the data. For example, Miozzo and Caramazza (1997b) found that participants in TOT states were better at freely producing partial phonological information than syntactic information (for a discussion of TOT experiments investigating access of syntax and phonology, see Biedermann, Ruh, Nickels & Coltheart, 2008). In addition, Miozzo and Caramazza (1997a) demonstrated a double dissociation between two Italian individuals with anomia: one participant was able to identify grammatical properties of a word he was unable to name. In contrast, another patient (originally presented by Badecker et al., 1995) was able to correctly produce phonological information orally (i.e., the correct word form) but not grammatical information (i.e., produced the wrong gender). Considered together these studies were used as evidence that syntax does not mediate between semantics and phonology but rather supported an architecture where phonological retrieval is not dependent on prior syntactic retrieval. This functional architecture makes a separate lemma level redundant

The Independent Network model also has a strictly feedforward architecture. However, unlike the Two-Stage model, activation may 'cascade' between levels: processing at one level need not be complete before it begins at the next level. In this model, frequency effects are located at the lexical (word form) level.

To summarise, in the Independent Network model activation is fed forward from semantics through a single (word form) level to the phoneme level. This word form level houses both syntactic and phonological networks that can be activated sequentially.

The Interactive Activation model

One of the earliest and pioneering models of word production is Dell's (1986) Interactive Activation model. The model has been subject to subsequent revisions and to date the most commonly cited version is Dell et al. (1997). It is this model that we focus on as it has been fully computationally implemented. However, an earlier model

(Dell, 1990) will also be outlined as this particularly focused on homophone processing.

Dell and colleagues' (1997) Interactive Activation model of lexical access assumes both feedforward and backward spreading activation. As Figure 4 shows, semantic features, word nodes and segments are arranged in layers with bidirectional links between them. The interactive activation flow is the crucial difference from Levelt et al.'s (1999) and Caramazza's et al. (2001) theoretical frameworks.

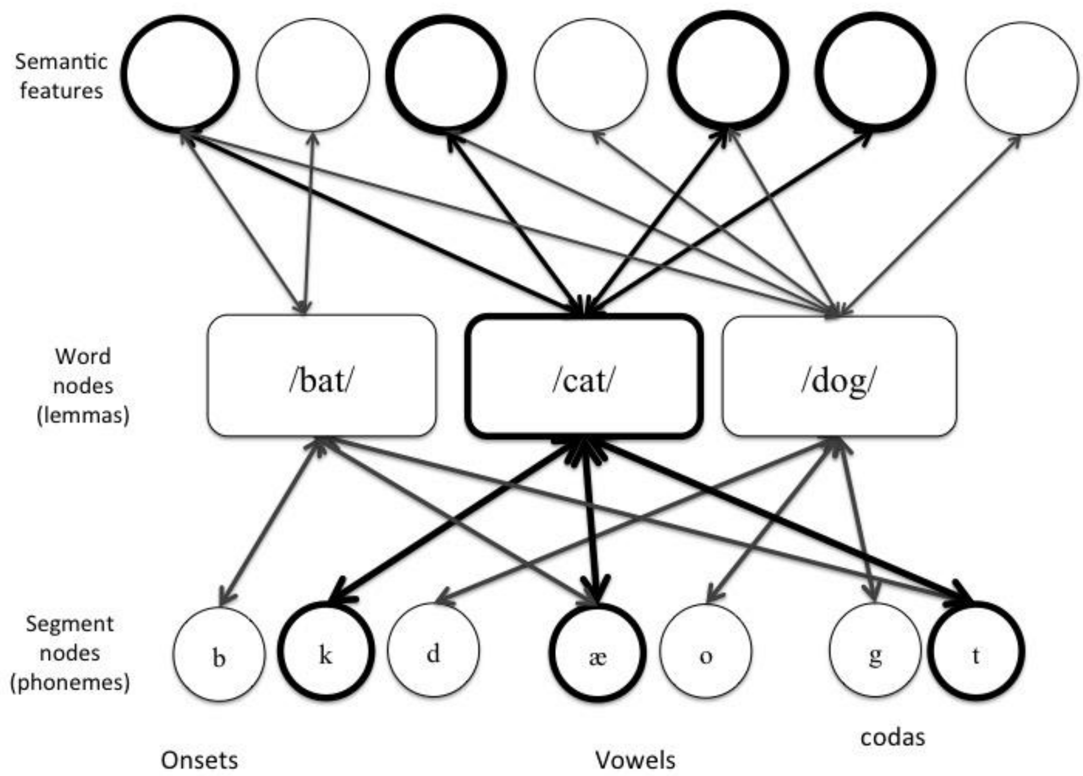
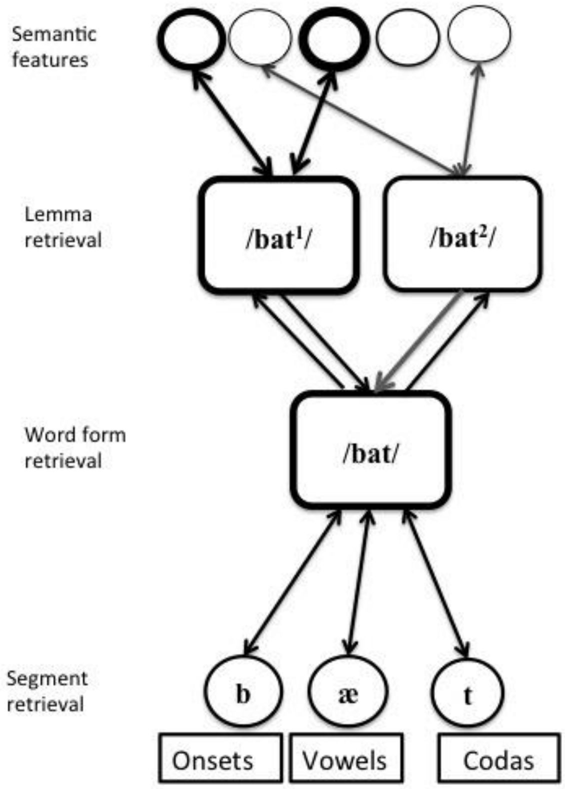


Figure 4. Representation of semantic features, words, and phonemes in the Interactive activation model (Dell et al., 1997).

Dell (1990, Figure 5) has a different architecture that also includes whole-word phonological representations that mediate between lemmas and phonemes, making it more similar to Levelt et al. (1999) in its overall architecture. However, the later instantiations of Dell's model do not include holistic phonological word form representations. Nevertheless, Schwartz et al. (2006) do suggest a more complete model that might specify lemma and phonological word form representations at separate levels. Instead, this information is currently represented in the links



between the lemma and phoneme levels (Schwartz et al., 2006) .

Figure 5. The representation of homophones in Dell's (1990) Interactive Activation model.

Word retrieval in Dell et al. (1997) is analogous to lemma access in Dell (1990). It involves retrieving a holistic, modality-neutral word node representation that is associated with grammatical properties (not explicitly represented). Lexical retrieval begins with a surge of activation to the semantic features of a target, this activation spreads bi-directionally throughout the model until the most active word is chosen for selection.

In most versions of this model, there is no competitive process (e.g., Dell et al., 1997), and the most active lexical node is selected after a predetermined number of time steps. However, in some versions of the Interactive Activation model, words compete for selection against each other (Dell, Lawler, Harris, & Gordon, 2004; Middleton, Chen, & Verkuilen, 2015) with selection being dependent on one item reaching a pre-determined activation threshold. Lexical competition is thought to arise post-semantically and pre-phonologically (Damian, Vigliocco, & Levelt, 2001), therefore is suggested to arise between nodes at the lemma level (e.g., Oppenheim, Dell, & Schwartz, 2010).

Regardless of whether due to competition or number of time steps, the most active word node is selected and subsequently receives a boost of activation. Once this word node is retrieved, the surge in activation spreads throughout the model increasing the activation of the target's phonemes. This results in the most active phonemes being selected for retrieval. As the boost of activation is much larger than any activation from feedback, the effect of interactivity is limited, making this model a compromise between discrete models (e.g., Caramazza et al., 2001; Levelt et al., 1999) and highly interactive models (e.g., Rapp & Goldrick, 2000). Due to the activation boosting mechanism, coupled with interactivity, the model predicts more semantic influence on word production earlier in the model- at the word retrieval stage. Nevertheless, unlike strictly feedforward models, the interactivity within the model results in phonological information affecting how semantics maps onto lexical representations (during word node retrieval). Therefore semantics is not the sole influence on word retrieval at the lemma stage as, due to interaction, phonology has some effect. The same is true during phonological retrieval: there is a greater influence of phonological overlap, with relatively limited, but not zero, influence of semantic overlap. This semantic influence resulting from interactivity provides a good explanation as to why mixed errors (semantic and phonological relationship to the target, e.g., *rat* for *cat*) may occur (Schwartz et al., 2006). In feedforward models these errors can only occur by chance. Their occurrence at rates greater than chance (Rapp & Goldrick, 2000) can only be explained in feedforward models through an external monitoring mechanism (e.g., Levelt, 1983, 1992) that is less likely to detect mixed errors.

Dell (1990) discusses a number of possibilities for the representation of frequency. One possibility is that frequency effects arise at the lemma level whereby higher frequency words could have a higher resting level of activation than lower frequency words (following Morton, 1969). However, he also explained how this is similar to the concept that higher frequency words have stronger connections from the semantic to the lemma level[2].

In brief, the Interactive Activation model encapsulates feedforward and feedback flow of activation between all adjacent levels. This interactivity results in

[2]Dell also considered a 'number of contexts' view, where frequency is represented by more connections/links to higher level nodes that represent the number of contexts in which that lemma has been used.

(some) influence of semantics and phonology on both word node and phoneme retrieval.

Representation of Homophones in the Lexicon

The Two-Stage, Interactive Activation, and Independent Network model all make specific predictions according to their functional architecture. Here, we discuss how these differences place unique constraints on homophone representation in each model.

Homophone representation in the Two-Stage model

Support for the Two-Stage model stems from a series of experiments investigating the locus of the frequency effect in the lexicon by Jescheniak and Levelt (1994). One of these experiments involved bilingual participants translating English words (e.g., *forest*) into low frequency Dutch homophones (e.g., *bos),* which also had a higher frequency homophone partner (e.g., *bos* means both forest and bunch in Dutch). It was found that the homophones were translated as fast as the summed frequency of both homophone meanings would predict, and significantly faster than the individual homophone frequency would predict. This homophone advantage (alongside a series of other experiments in Jescheniak & Levelt, 1994) led Levelt et al. (1999) to conclude not only that frequency was located at word form level, but also that homophones shared word form representations and hence inherit the frequency of their homophone partner (as shown in Figure 6A). Further support that homophones have separate lemmas but shared phonological word forms has been found in both psycholinguistic and neuropsychological studies (Biedermann, Blanken,

& Nickels, 2002; Biedermann & Nickels, 2008b, 2008a; Biran, Gvion, Sharabi, & Gil, 2013; Jescheniak, Meyer, & Levelt, 2003).

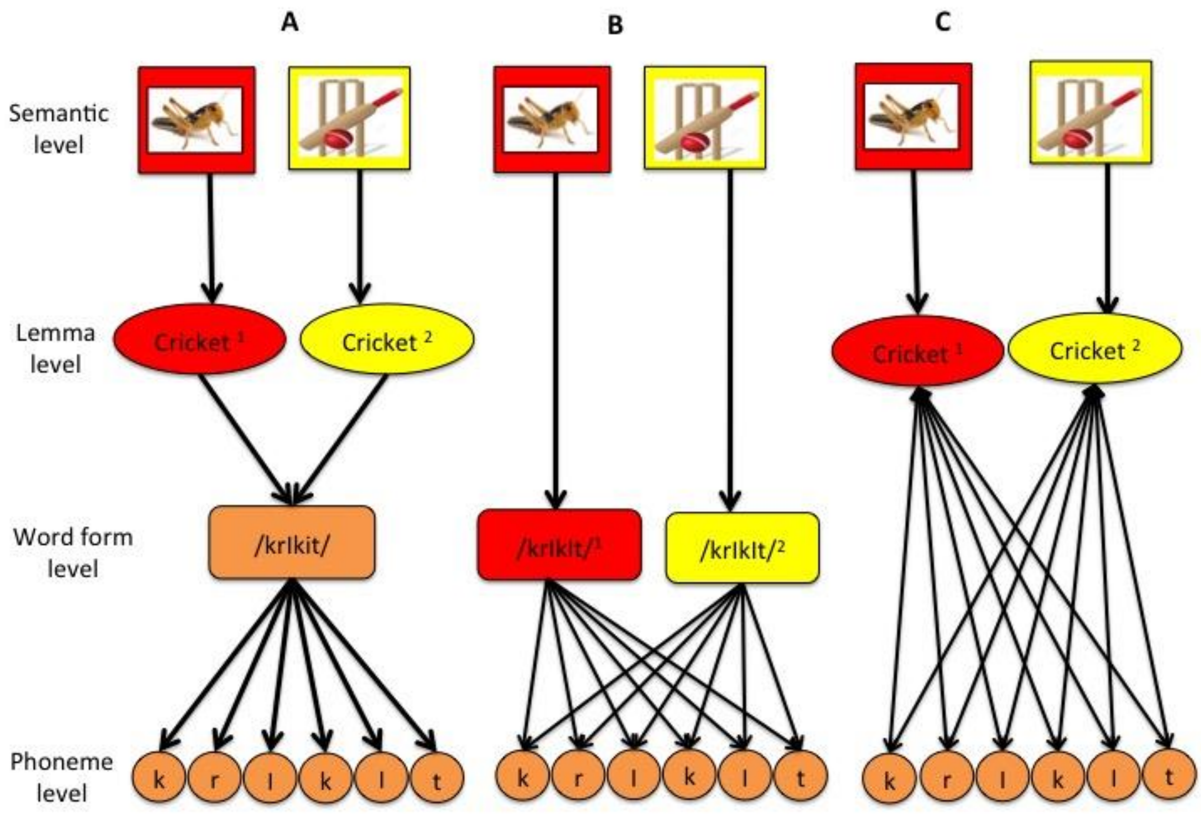


Figure 6. Homophone representation in A) Two-Stage model (Levelt et al., 1999) B) the Independent Network model (Caramazza et al., 2001) and C) the dual nature account (Middleton et al., 2015).

Homophone representation in the Independent Network model

Caramazza et al. (2001) failed to replicate the homophone translation effect found in Jescheniak and Levelt's (1994) experiment. In their experiments, English homophones were translated from Spanish words at latencies predicted by their individual frequency and significantly slower than their summed frequency. Caramazza et al., 2001 also investigated the homophone effect in picture naming and found the same pattern of results: homophone naming latencies were according to individual frequency. Caramazza et al. (2001) concluded therefore that homophones are produced according to their individual word frequency and therefore are represented independently as shown in Figure 6B.

Homophone representation in the Interactive Activation model

The 1990 version of Dell's Interactive Activation model (Dell, 1990; Figure 5, earlier) includes separate lemma and word form levels and specifically addresses representation of homophones (which is not the case in later models). Dell (1990) found that low frequency homophones (e.g., hymn) were produced as accurately as their high frequency partners (e.g., him), which draws parallels with Jescheniak and Levelt's (1994) finding of a homophone advantage in translation tasks. Like Levelt et al. (1999), Dell (1990) postulated that homophones share a representation at the lexeme/word form level. Uniquely, due to activation of the shared word form, coupled with the interactive processes of the Interactive Activation model, activation feeds back to the higher frequency (non-target) homophone lemma. This, in turn, feeds activation down to the shared word node resulting in the lower frequency homophone inheriting the higher frequency of the homophone partner (despite frequency being located at the lemma and not the word form level).

However, more recent versions of the Interactive Activation model have not included two separate lexical levels. Instead, the current version of the model contains a single lemma level that links to both lexical-syntax and represents the phonological word form in the links to the phoneme level (Figure 6c; e.g., Dell et al., 2004, 1997; Schwartz et al., 2006). Consequently, within this architecture, homophones are, necessarily, represented as independent nodes.

Middleton et al. (2015) explicitly discuss homophone representation in terms of a version of the Interactive Activation model which includes competition at the lemma level (e.g., Dell et al., 2004; Schwartz et al., 2006). Their Dual Nature account of homophone processing is supported by evidence from aphasia. Middleton et al. (2015) contrast the homophone naming of participants with aphasia who presented with semantic impairments (Stage-1 Impairments, analogous to word node (lemma) impairments in e.g., Schwartz et al., 2006) and those who have phonological deficits (Stage-2 deficits). While for the Stage 1 impairment group, a homophone naming disadvantage was found, Stage-2 participants showed a homophone advantage. In the Dual Nature account, due to the interactive and competitive nature of the model, Middleton et al. (2015) suggest that there are two possible factors influencing homophone production in spoken production. First, during Stage-1 (lemma) retrieval, homophones can have a disadvantage due to feedback from the shared phonemes to two different lemmas (Middleton et al., 2015). This feedback results in greater competition than is the case for non-homophones where other lemmas would be less active due to fewer overlapping phonemes. In contrast, during Stage-2 retrieval, there is an advantageous effect of homophony on phoneme selection because of reverberated activation from the second lemma node with the same phonology.

Therefore, the most current representation of homophones within the Interactive Activation model by Dell and colleagues is Middleton et al.'s (2015) Dual Nature account model as depicted in Figure 4 From here on, I will refer to Middleton et al.'s model (2015) when discussing homophone representation within the Interactive Activation model.

In summary, in this model homophones have separate lemma nodes (which is where frequency effects originate). Yet, due to the interactive nature of the model, and competition at the lemma level, homophony can be both advantageous for Stage-2 retrieval and disadvantageous for Stage-1 retrieval in production. Hence, the precise pattern observed may depend on the balance between these factors.

Written production

There has been considerably less research into written production, particularly written homophone production. Before I will go into the intricacies of homophone representation(s) in the written lexicon, I will outline how a written word is produced.

It is fairly well established that there are several routes available to spell a word, as shown in Figure 7 (e.g., Bonin, Peereman, & Fayol, 2001; Houghton & Zorzi, 2003; Tainturier & Rapp, 2001). The use of each pathway depends on the task at hand and whether the target is familiar (word) or novel (nonword). If an item is familiar, then to ensure accurate spelling, a stored representation of the spelling must be 'looked up' in the orthographic output lexicon. Nonwords (e.g., clup) cannot be spelled via this route and instead are spelled using a rule-governed procedure (labelled Route a in Figure 7).

In written picture naming, first a concept will be activated in the semantic system, this activation then spreads to the orthographic lexicon, where an abstract word form is activated which in turn activates the letters of the word in the graphemic buffer for production (by hand, typing or oral spelling). Although underspecified, it is implicitly suggested that the lexical links between the modality-neutral semantic system and the orthographic word form level are feedforward only (Bonin & Fayol, 2000). Items with irregular spelling (i.e. words which are not spelled phonetically-such as *yacht*) can only be produced accurately via Route b (Figure 7).

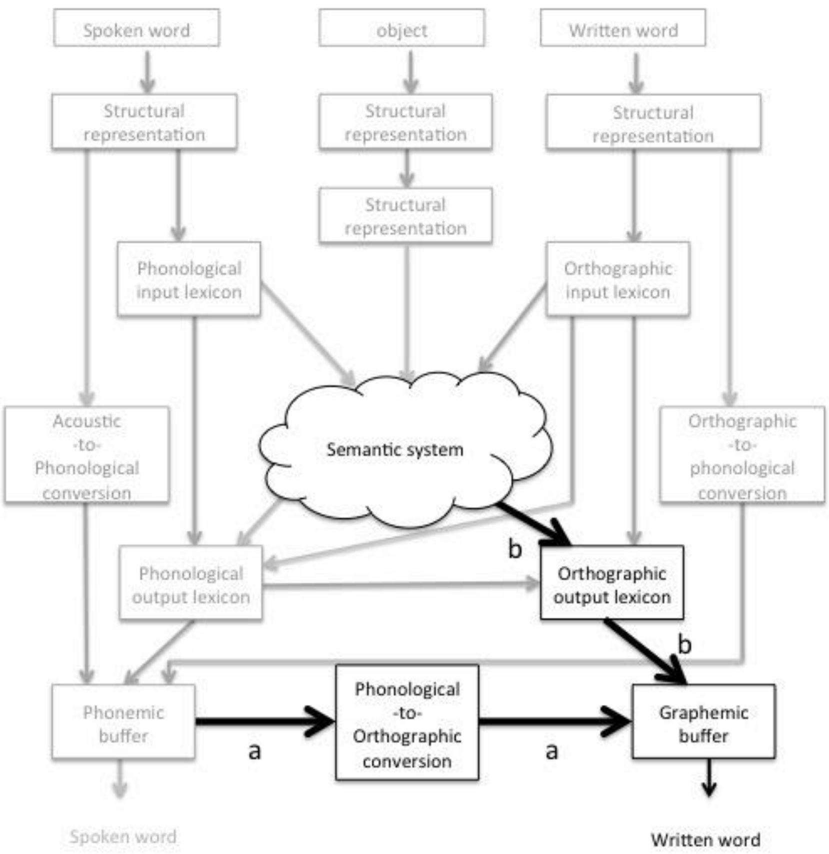


Figure 7. A language production model focusing on spelling. The bold links show the most activated, spelling-specific components for a) writing nonwords to dictation and b) written picture naming on the basis of Bonin, Méot, Lagarrigue, and Roux (2015).

The other, non-lexical, route is utilised for spelling unfamiliar words or non-words. When an unknown word is presented for spelling, the phonemes (activated in the phonemic buffer) are converted into their corresponding graphemes and computed as a word via the non-lexical route. The output for regular (and consistent) words will most likely be correct, whereas the result for an irregular word will be incorrect via this route.

Both routes for spelling lead to the graphemic buffer where the abstract letter sequences are kept active for further processing. As in spoken word production, activation may be feedforward (Caramazza, 1997) or there may be feedback from the graphemic buffer to the lexicon (e.g., McCloskey, Macaruso, & Rapp, 2006).

The type of task will also influence the route taken to spell an item. As can be inferred from Figure 7, written picture naming, writing to dictation and copying all have the potential to access the semantic system. However, it has been found that while they all can potentially access semantic representations, the extent of semantic and/or lexical activation differs across tasks (Bonin et al., 2015). Written picture naming requires the most access to the orthographic lexicon (lexical route), whereas only writing to dictation reliably accesses the sub-lexical route.

The majority of writing research has involved writing to dictation, which need not rely on accessing orthographic representations from semantics (Bonin et al., 2015). As we are predominantly interested in orthographic lexical representations and how they are accessed (as this is the level which may or may not be shared for homophones), we focus on the process of written picture naming (which does rely on the lexical route, Bonin et al., 2015). Due to evidence from impaired spellers (e.g. Rapp & Caramazza, 1997) and unimpaired spellers (Bonin, Fayol, & Peereman, 1998), it has largely been rejected that written naming must be mediated via the phonological output lexicon. Instead, as noted above, it is assumed that the orthographic output lexicon can be directly activated from the semantic system (without prior phonological access; Bonin, Chalard, Méot, & Fayol, 2002; Tainturier & Rapp, 2001). Nonetheless, it is plausible that phonology will also be activated in spelling (just not obligatorily mediated) as it has been found to influence spelling

(Bonin et al., 2001), most likely via links between the phonological and orthographic output lexicons (e.g., Bonin et al., 2001; Patterson, 1986).

Comparing the Phonological and Orthographic Output Lexicons

As mentioned above, there is general consensus that the phonological and orthographic output lexicons are distinct. However, it is not known how similar these two distinct systems are in their representations. Few studies have investigated the similarity between spoken and written production. However, Bonin et al. (2002) investigated the effect of numerous variables on spoken and written picture naming latencies. They found that age of acquisition, image variability, name agreement and image agreement affect both types of picture naming. The similarity of the effects in both modalities is a reasonable basis on which to conclude that there are some resemblances across modalities (Bonin & Fayol, 2000).

One of the few studies that has explicitly compared spoken and written production investigated homophone picture naming. Bonin and Fayol (2002) found that high frequency (heterographic) homophones were produced faster than their low frequency counterparts in both written and spoken modalities (despite written production having overall longer latencies). This again supports the theory that the processes and representations may be similar across the two modalities.

Figure 8 depicts the only clear specification of both phonological and orthographic homophone representation (Caramazza & Miozzo, 1997). Earlier work by Caramazza and Hillis (1991) presented two participants with aphasia (difficulty producing words after brain damage) who had opposite impairments. Both participants were impaired when producing homophone verbs compared to nouns (e.g. impaired in in producing 'to *watch'* but intact production of 'the *watch'*). Critically, however, one patient was impaired in producing the verb in oral production but unimpared in written production, whereas the opposite was true for the other participant. This double dissociation supported a model where homophones (at least cross class homographs) are represented independently in each lexicon (Caramazza & Hillis, 1991; Hillis & Caramazza, 1995; Rapp & Caramazza, 1997).

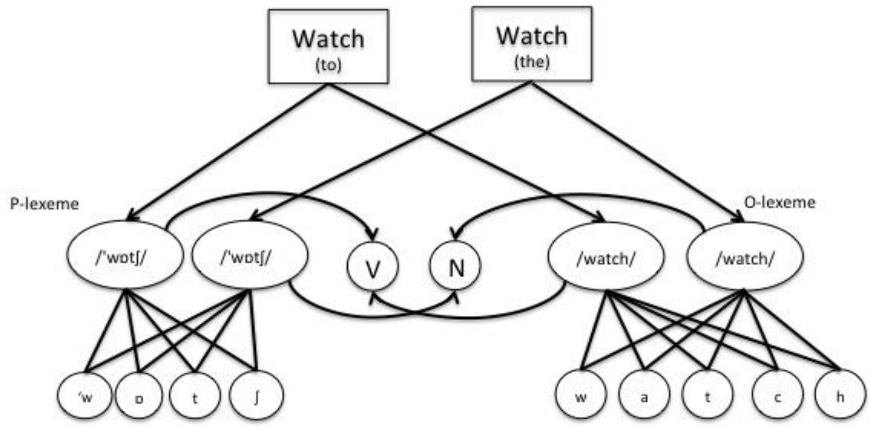


Figure 8. Cross class homograph representation in the orthographic and phonological lexicons, adapted from Caramazza (1997).

Although it is not known how homophones are represented in the orthographic output lexicon, this research, along with work of Bonin and colleagues, suggests that there is the potential for homophones to be represented similarly across lexicons (at least for homographic homophones).

Methods used to Investigate Homophone Representation

A variety of different methodologies have previously been used in research regarding the nature of homophone representations, here we briefly describe these methods - further detail on the results of these studies and the conclusions drawn is provided, as appropriate, in the experimental chapters of the thesis.

Frequency inheritance

Dell (1990) introduced the term frequency inheritance as he found low frequency homophones 'inherited' protection from errors from their high frequency counterparts. Subsequent studies investigating this phenomenon have therefore compared the production of homophones to control words matched to either the individual frequency of one of the homophone pair or the summed frequency of both homophone meanings. If there is frequency inheritance, then homophones should behave similarly to their summed controls, resulting in, for example, production equivalent to the speed (Jescheniak & Levelt, 1994) or accuracy (Middleton et al., 2015) of summed frequency controls. However, if there is not a homophone frequency inheritance, then homophones will be produced as fast (e.g. Caramazza et al., 2001) or as accurate (e.g., Jacobs, Singer, & Miozzo, 2004, although see Middleton et al., 2015 for a review of this paper) as individual frequency matched controls.

Bonin and Fayol (2002) investigated frequency inheritance in a spoken and written heterographic homophone picture naming task. The authors found a lack of frequency inheritance in that low frequency homophones were produced significantly slower than high frequency homophones.

Although the frequency inheritance paradigm is a useful method to investigate representations within the lexicon, the conclusions that can be drawn from this paradigm critically rest on assumptions regarding the location of frequency in the models. Hence, other paradigms have been used that do not rely on this contentious issue.

Aphasia treatment/ generalisation

Aphasia is an acquired language disorder resulting from brain damage. The most common impairments in aphasia are word retrieval deficits (e.g., Dell et al., 1997). Successful phonological treatment of aphasia via repeatedly producing an item is hypothesied to be due to strengthened links between representations (e.g. to or from the word form; Nickels, 2002; Biedermann & Nickels, 2008a, 2008b) improving word retrieval. This is the same mechanism that has been hypothesised to be behind long-term priming (Cave, 1997).

Treatment of aphasic naming has been used to explore homophone representations specifically because it circumvents the debate of where frequency is located. Blanken (1989) developed the concept of investigating homophone representation by examining generalisation of aphasia therapy effects, which was followed by a series of studies by Biedermann and colleagues aimed at improving phonological retrieval in people with phonological output impairments in aphasia (Biedermann et al., 2002; Biedermann & Nickels, 2008a, 2008b). They treated one homophone (e.g. *ball*- the dance) to examine whether there would be treatment-related improvement of the untreated partner (e.g. *ball*-the toy) (Biedermann et al., 2002; Biedermann & Nickels, 2008b, 2008a; Biran et al., 2013). The reasoning behind this methodology is that if treatment causes improvement to untreated items (generalisation) this implies that, at some level, the items that show generalisation share, or overlap in, their representation, or are functionally linked to the treated items (Nickels et al., 2010). Therefore, generalisation is a useful method for investigating whether homophones share representations at the word form level (resulting in generalisation through improved access to the single representation), have independent representations (no generalisation, with improved access to only the treated representation), or interactive feedback (generalisation dependent on overlap from segments to word nodes, improved access to the treated representation and feedback during treatment to the untreated representation also improving its access).

Priming

A robust finding is that prior production of the *same* item will result in faster subsequent production time, compared to prior production of an *unrelated* word (e.g., Barry, Hirsh, Johnston, & Williams, 2001; Ferrand, Grainger, & Segui, 1994; Wheeldon & Monsell, 1992). This 'repetition' priming effect has been found immediately after presentation of prime and with long durations between prime and target (e.g., Cave, 1997), as well as within and across different modalities (Damian, Dorjee, & Stadthagen-Gonzalez, 2011).

The most prominent theory of priming mechanisms implicates ‘structural’ changes in the accessibility of lexical representations (e.g., Barry et al., 2001). Specifically, previous processing of a lexical entry can change the activation levels required to retrieve the stored representations. Wheeldon and Monsell (1992) point out that this could be for a number of reasons: either due to lowering of the activation thresholds required for an item to be chosen, or that there is a change in the resting activation state of the entries or due to a changing of the weights of connections between levels. Regardless of the exact mechanism, structural change results in faster production for a previously produced item.

Wheeldon and Monsell (1992) investigated whether priming requires previous activation of semantics, as well as, word form, by priming homophone naming with previous naming to definition of their partners. They found no robust priming effect (although there was some evidence of homograph priming but this was not replicated across analyses), concluding that priming required top down (i.e., semantic) activation to induce structural change resulting in subsequent facilitation. However, Cutting and Ferreira (1999) found that a written word that was semantically related to the homophone partner (e.g. for the homophone *ball* (the social event) the word dance) caused facilitation of subsequent picture naming of target homophone (e.g. *ball*- the toy). Furthermore, in another study, semantically priming a homophone noun partner resulted in incorrectly writing a noun inflection (e.g. 's') on a homophonic verb (Largy, Fayol, & Lemaire, 1996). Similarly, White, Abrams, McWhite and Hagler (2010) found increased homophone spelling errors in writing sentences to dictation (e.g. writing *beech* as beach) when the homophone partner had been primed by an orthographically and/or phonologically related prime (e.g. *teacher*)

These oblique homophone priming findings (Cutting & Ferreira, 1999; Largy et al., 1996) suggest that even indirect previous activation of a homophone can subsequently affect production of the target. Crucially, however, there are no published attempts to replicate Wheeldon and Monsell's (1992) priming study. Intriguingly, this study did provide some suggestion that priming may be affected by homophone spelling, as an inhibitory effect was found for heterographic priming compared to a small facilitatory effect of homograph priming. Replication and extension of this research would be an important addition to the literature.

Questions Left to Answer

Despite the not insubstantial body of research on homophones outlined above, there are still questions that are left to be answered in order disentangle homophone representation.

Firstly, how are homophones represented in the phonological lexicon? The following three options remain possible: (a) shared at the word form level (Two-Stage model; Levelt et al., 1999), (b) independent at the word form level with strictly feedforward activation from word forms to segments (Independent Network model; Caramazza et al., 2001) or (c) independent word form representations with interactive activation flow between segments and word forms (the dual nature account; Middleton et al., 2015). These accounts lead to the empirical question of whether homophones have a production advantage, no advantage or a pattern that varies according to the nature of the task. In addition, it remains to be determined whether the same pattern of effects and nature of representations are found in the orthographic lexicon as in the phonological lexicon. Finally, there is the question of whether homophone representation differs depending on their spelling and whether the resulting impact on processing interacts with the modality of output?

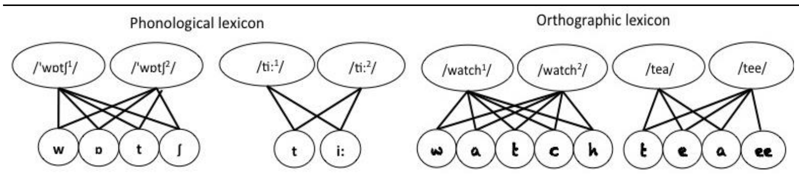
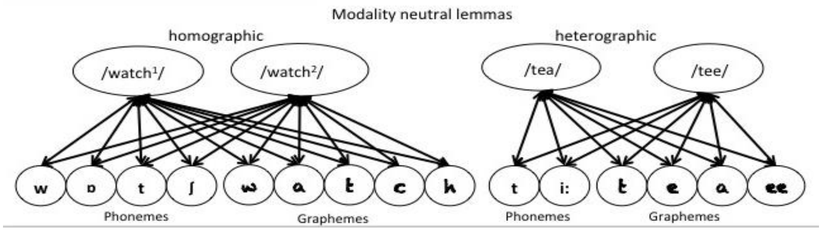
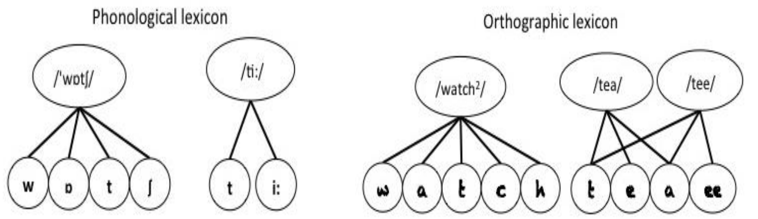
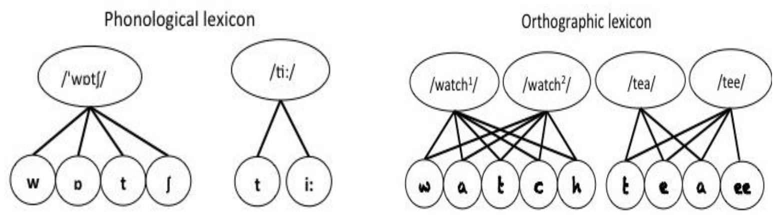
Table 1 depicts the various possible architectures of the phonological and orthographic lexicon in regards to homophone representation (column 1). The subsequent columns show the prediction for whether homophones should show a processing advantage compared to non-homophones (over and above any advantage predicted by the overlap of segments). By conducting research using a combination of methodologies and comparing their outcomes, this thesis should result in a more comprehensive understanding of homophone representation(s) and help to discriminate between these theoretical possibilities.

Preview of Thesis

This thesis aims to address the questions raised above using a range of methodologies and participant populations.

Chapter 2 investigates spoken homophone production in two experiments examining picture naming and translation. This study aims to examine homophone representation and processing in spoken word production and specifically focuses on potential reasons why the past literature in this domain may have produced contradictory results.

Table 1. The possible representations of homographs and heterographs in the phonological and orthographic lexicon and their predicted effects.

Representation	**Predicted Advantage?**			
	Spoken		**Written**	
	Homo	Het	Homo	Het
Independent in both lexicons and feedforward activation e.g. Independent Network	✗	✗	✗	✗
Independent in both lexicons with feedback and competition e.g. Dual nature account	✓/✗	✓/✗	✓/✗	?
Shared in phonological lexicon, only homographs shared in orthographic and feedforward activation e.g., The most likely version of the Two-Stage model adapted to include the orthographic lexicon	✓	✓	✓	✗
Shared in phonological lexicon only and feedforward activation e.g. an alternative adaptation of the Two-Stage model to include the orthographic lexicon	✓	✓	✗	✗

Homo= homograph
Het= heterograph
✓ Advantage e.g. frequency inheritance, generalisation of homophone treatment, homophone priming
✗ No advantage e.g. homophones produced as individual frequency would suggest, no generalisation and priming or only to the extent predicted by segment overlap/feedback,
✓/✗ Both advantage and no advantage depending on the stage of retrieval

Chapter 3 uses a different methodology – generalisation of treatment effects in aphasia to inform our understanding of orthographic representation of homophones.

The final experimental chapter, Chapter 4, uses a fourth methodology – priming of written picture naming- to also inform our understanding of homophone representation and processing in the orthographic modality.

Comparing the findings from these three experimental studies will not only inform theories of homophone representation across orthographic and phonological lexicons, but will also inform us how similar these lexicons are to each other. By comparing heterographs (*flour* vs. *flower*) to homographs (*ball* vs. *ball*) these experiments will enable investigation of the extent to which orthography influences phonology (Chapter 1) and phonology influences orthography (Chapters 2 and 3). Replicating findings across different studies will enhance validity (e.g., Muma, 1993) and will help to answer open questions in regards to homophone production. Additionally, this series of experiments should further constrain which of the homophone representations depicted in Table 1 is more plausible. Ultimately, defining homophone representation(s) and their effects on word production will inform more general theories of language processing (i.e. the flow of activation, the number of levels in lexical retrieval, and the internal organisation at each level, etc.).

References

Badecker, W., Miozzo, M., & Zanuttini, R. (1995). The Two-Stage model of lexical retrieval: evidence from a case of anomia with selective preservation of grammatical gender. *Cognition*, *57*(2), 193–216. https://doi.org/10.1016/0010-0277(95)00663-J

Barry, C., Hirsh, K. W., Johnston, R. A., & Williams, C. L. (2001). Age of Acquisition, Word Frequency, and the Locus of Repetition Priming of Picture Naming. *Journal of Memory and Language*, *44*(3), 350–375. https://doi.org/10.1006/jmla.2000.2743

Biedermann, B., Blanken, G., & Nickels, L. (2002). The representation of homophones: Evidence from remediation. *Aphasiology*, *16*(10–11), 1115–1136. https://doi.org/10.1080/02687030244000545

Biedermann, B., & Nickels, L. (2008a). Homographic and heterographic homophones in speech production: Does orthography matter? *Cortex*, *44*(6), 683–697. https://doi.org/10.1016/j.cortex.2006.12.001

Biedermann, B., & Nickels, L. (2008b). The representation of homophones: More evidence from the remediation of anomia. *Cortex*, *44*(3), 276–293. https://doi.org/10.1016/j.cortex.2006.07.004

Biedermann, B., Ruh, N., Nickels, L., & Coltheart, M. (2008). Information Retrieval in Tip of the Tongue States: New Data and Methodological Advances. *Journal of Psycholinguistic Research*, *37*(3), 171–198. https://doi.org/10.1007/s10936-007-9065-8

Biran, M., Gvion, A., Sharabi, L., & Gil, M. (2013). The Representation of Homophones in the Phonological Lexicon: Additional Evidence from Hebrew. *Procedia - Social and Behavioral Sciences*, *94*, 86–87. https://doi.org/10.1016/j.sbspro.2013.09.041

Blanken, G. (1989). Wortfindungsstorungen und Verbales Lernen bei Aphasie- Eine Einzelfallstudie. *Neurolinguistik*, 2, 107-126.

Bonin, P., Chalard, M., Méot, A., & Fayol, M. (2002). The determinants of spoken and written picture naming latencies. *British Journal of Psychology*, *93*(1), 89–114. https://doi.org/10.1348/000712602162463

Bonin, P., & Fayol, M. (2000). Writing words from pictures: What representations are activated, and when? *Memory & Cognition*, *28*(4), 677–689. https://doi.org/10.3758/BF03201257

Bonin, P., & Fayol, M. (2002). Frequency effects in the written and spoken production of homophonic picture names. *European Journal of Cognitive Psychology*, *14*(3), 289–313. https://doi.org/10.1080/09541440143000078

Bonin, P., Fayol, M., & Peereman, R. (1998). Masked form priming in writing words from pictures: Evidence for direct retrieval of orthographic codes. *Acta Psychologica*, *99*(3), 311–328. https://doi.org/10.1016/S0001-6918(98)00017-1

Bonin, P., Méot, A., Lagarrigue, A., & Roux, S. (2015). Written object naming, spelling to dictation, and immediate copying: Different tasks, different pathways? *The Quarterly Journal of Experimental Psychology*, *68*(7), 1268–1294. https://doi.org/10.1080/17470218.2014.978877

Bonin, P., Peereman, R., & Fayol, M. (2001). Do Phonological Codes Constrain the Selection of Orthographic Codes in Written Picture Naming? *Journal of Memory and Language*, *45*(4), 688–720. https://doi.org/10.1006/jmla.2000.2786

Caramazza, A. (1997). How Many Levels of Processing Are There in Lexical Access? *Cognitive Neuropsychology*, *14*(1), 177–208. https://doi.org/10.1080/026432997381664

Caramazza, A., Costa, A., Miozzo, M., & Bi, Y. (2001). The specific-word frequency effect: Implications for the representation of homophones in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(6), 1430–1450. https://doi.org/10.1037/0278-7393.27.6.1430

Caramazza, A., & Hillis, A. E. (1991). Lexical organization of nouns and verbs in the brain. *Nature*, *349*(6312), 788–790. https://doi.org/10.1038/349788a0

Caramazza, A., & Miozzo, M. (1997). The relation between syntactic and phonological knowledge in lexical access: evidence from the `tip-of-the-tongue' phenomenon. *Cognition*, *64*(3), 309–343. https://doi.org/10.1016/S0010-0277(97)00031-0

Cave, C. B. (1997). Very Long-Lasting Priming in Picture Naming. *Psychological Science*, *8*(4), 322–325. https://doi.org/10.1111/j.1467-9280.1997.tb00446.x

Cutting, J., & Ferreira, V. (1999). Semantic and phonological information flow in the production lexicon. *J Exp Psychol Learn Mem Cogn*, *25*(2), 318–344.

Damian, M. F., Dorjee, D., & Stadthagen-Gonzalez, H. (2011). Long-term repetition priming in spoken and written word production: Evidence for a contribution of phonology to handwriting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(4), 813–826. https://doi.org/10.1037/a0023260

Damian, M. F., Vigliocco, G., & Levelt, W. J. . (2001). Effects of semantic context in the naming of pictures and words. *Cognition*, *81*(3), B77–B86. https://doi.org/10.1016/S0010-0277(01)00135-4

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283–321. https://doi.org/10.1037/0033-295X.93.3.283

Dell, G. S. (1990). Effects of Frequency and Vocabulary Type on Phonological Speech Errors. *Language and Cognitive Processes*, *5*(4), 313–349. https://doi.org/10.1080/01690969008407066

Dell, G. S., Chang, F., & Griffin, Z. M. (1999). Connectionist models of language production: lexical access and grammatical encoding. *Cognitive Science*, *23*(4), 517–542. https://doi.org/10.1016/S0364-0213(99)00014-2

Dell, G. S., Lawler, E. N., Harris, H. D., & Gordon, J. K. (2004). Models of errors of omission in aphasic naming. *Cognitive Neuropsychology*, *21*(2–4), 125–145. https://doi.org/10.1080/02643290342000320

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, *104*(4), 801–838. https://doi.org/10.1037/0033-295X.104.4.801

Ferrand, L., Grainger, J., & Segui, J. (1994). A study of masked form priming in picture and word naming. *Memory & Cognition*, *22*(4), 431–441. https://doi.org/10.3758/BF03200868

Hillis, A. E., & Caramazza, A. (1995). Representation of Grammatical Categories of Words in the Brain. *Journal of Cognitive Neuroscience*, *7*(3), 396–407. https://doi.org/10.1162/jocn.1995.7.3.396

Houghton, G., & Zorzi, M. (2003). Normal and Impaired Spelling in a Connectionist Dual-Route Architecture. *Cognitive Neuropsychology*, *20*(2), 115–162. https://doi.org/10.1080/02643290242000871

Jacobs, M., Singer, N., & Miozzo, M. (2004). The representation of homophones: Evidence from anomia. *Cognitive Neuropsychology*, *21*(8), 840–866. https://doi.org/10.1080/02643290342000573

Jescheniak, J. D., & Levelt, M., J. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 824–843. https://doi.org/10.1037/0278-7393.20.4.824

Jescheniak, J. D., Meyer, A. S., & Levelt, W. J. M. (2003). Specific-word frequency is not all that counts in speech production: Comments on Caramazza, Costa, et al. (2001) and new experimental data. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(3), 432–438. https://doi.org/10.1037/0278-7393.29.3.432

Jones, A. C., & Rawson, K. A. (2016). Do reading and spelling share a lexicon? *Cognitive Psychology*, *86*, 152–184. https://doi.org/10.1016/j.cogpsych.2016.02.003

Largy, P., Fayol, M., & Lemaire, P. (1996). The Homophone Effect in Written French: The Case of Verb-Noun Inflection Errors. *Language and Cognitive Processes*, *11*(3), 217–256. https://doi.org/10.1080/016909696387178

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition*, *14*(1), 41–104. https://doi.org/10.1016/0010-0277(83)90026-4

Levelt, W. J. M. (1992). The perceptual loop theory not disconfirmed: A reply to MacKay. *Consciousness and Cognition*, *1*(3), 226–230. https://doi.org/10.1016/1053-8100(92)90062-F

Levelt, W. J. M. (1999). Models of word production. *Trends in Cognitive Sciences*, *3*(6), 223–232. https://doi.org/10.1016/S1364-6613(99)01319-4

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(01), 1–38.

McCloskey, M., Macaruso, P., & Rapp, B. (2006). Grapheme-to-lexeme feedback in the spelling system: Evidence from a dysgraphic patient. *Cognitive Neuropsychology*, *23*(2), 278–307. https://doi.org/10.1080/02643290442000518

Middleton, E. L., Chen, Q., & Verkuilen, J. (2015). Friends and foes in the lexicon: Homophone naming in aphasia. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(1), 77–94. https://doi.org/10.1037/a0037778

Miozzo, M., & Caramazza, A. (1997). Retrieval of lexical–syntactic features in tip-of-the tongue states. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(6), 1410–1423. https://doi.org/10.1037/0278-7393.23.6.1410

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, *76*(2), 165–178. https://doi.org/10.1037/h0027366

Muma, J. R. (1993). The need for replication. *Journal of Speech and Hearing Research*, *36*(5), 927–930.

Nickels, L. (2002). Therapy for naming disorders: Revisiting, revising, and reviewing. *Aphasiology*, *16*(10-11), 935-979.

Nickels, L., Kohnen, S., & Biedermann, B. (2010). An untapped resource: Treatment as a tool for revealing the nature of cognitive processes. *Cognitive Neuropsychology*, *27*(7), 539–562. https://doi.org/10.1080/02643294.2011.609811

Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition*, *114*(2), 227–252. https://doi.org/10.1016/j.cognition.2009.09.007

Patterson, K. (1986). Lexical but nonsemantic spelling? *Cognitive Neuropsychology*, *3*(3), 341–367. https://doi.org/10.1080/02643298608253363

Rapp, B., & Caramazza, A. (1997). The Modality-Specific Organization of Grammatical Categories: Evidence from Impaired Spoken and Written Sentence Production. *Brain and Language*, *56*(2), 248–286. https://doi.org/10.1006/brln.1997.1735

Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review*, *107*(3), 460.

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, *64*(3), 249–284.

Roelofs, A., Meyer, A., & Levelt, W. (1998). A case for the lemma/lexeme distinction in models of speaking: comment on Caramazza and Miozzo (1997). *Cognition*, *69*(2), 219–230.

Schwartz, M. F., Dell, G. S., Martin, N., Gahl, S., & Sobel, P. (2006). A case-series test of the interactive two-step model of lexical access: Evidence from picture naming. *Journal of Memory and Language*, *54*(2), 228–264. https://doi.org/10.1016/j.jml.2005.10.001

Tainturier, M.-J., & Rapp, B. (2001). The Spelling Process. In *The Handbook of Cognitive Neuropsychology: What Deficits Reveal about the Human Mind*. Psychology Press.

Vigliocco, G., Antonini, T., & Garrett, M. F. (1997). Grammatical Gender Is on the Tip of Italian Tongues. *Psychological Science*, *8*(4), 314–317. https://doi.org/10.1111/j.1467-9280.1997.tb00444.x

Wheeldon, L. R., & Monsell, S. (1992). The Locus of Repetition Priming of Spoken Word Production. *The Quarterly Journal of Experimental Psychology Section A*, *44*(4), 723–761. https://doi.org/10.1080/14640749208401307

White, K. K., Abrams, L., McWhite, C. B., & Hagler, H. L. (2010). Syntactic constraints in the retrieval of homophone orthography. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *36*(1), 160–169. https://doi.org/10.1037/a0017676

Wilshire, C. E. (2008). Cognitive neuropsychological approaches to word production in aphasia: Beyond boxes and arrows. *Aphasiology*, *22*(10), 1019–1053. https://doi.org/10.1080/02687030701536016

Wilshire, C. E., & McCarthy, R. A. (2002). Evidence for a context-sensitive word retrieval disorder in a case of nonfluent aphasia. *Cognitive Neuropsychology*, *19*(2), 165–186.

Chapter 2

Two bee or not to be: The Effects of Orthography and Bilingualism on Spoken Homophone Production

Theories of speech production all postulate that in order to produce a word, first, an abstract semantic (meaning) representation is activated and then grammatical and phonological knowledge is retrieved (lexical retrieval). Some theories (such as the Two-Stage model, Levelt, Roelofs, & Meyer, 1999) propose two separate levels activated sequentially in lexical retrieval; the first representing grammatical knowledge (at the lemma level; e.g., number, grammatical gender, word class) and then phonological knowledge (at the word form level). Other theories represent grammatical and phonological information at the same level (e.g., Independent Network model; Caramazza, Costa, Miozzo, & Bi, 2001) or a single lexical level with phonological information being represented in the links leading from this level to phonemes (the Interactive Activation model; Schwartz, Dell, Martin, Gahl, & Sobel, 2006).

Homophones are words that have two separate meanings but one pronunciation. Homophones can either be spelled the same (homographic), such as *cricket*- the insect and *cricket*-the game, or spelled differently (heterographic), such as *knight* and *night*. The fact homophones have separate meanings but shared phonology means they are a useful tool for investigating the processing steps from semantics to phonology. Currently, there are three theories regarding the representation of homophone lexical retrieval. All theories posit that homophones have separate conceptual/semantic representations for each meaning, however, after this level theories diverge: (i) homophones have independent lexical-syntactic representations (lemmas) but share a representation at the word form level (as depictured in the Two-Stage model, Levelt et al., 1999, shown in Figure 1A), (ii) there is no intervening lexical-syntactic level and homophone words have separate word forms (which link to lexical-syntax) for each meaning (as suggested by the Independent Network model; Caramazza et al.,, 2001, see Figure 1B), or (iii) homophones have independent lemmas and the links from the lemmas to a phoneme level represent the phonological form and, critically, these links are interactive with activation flowing from lemma to phonemes and also back from phonemes to lemma (Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Middleton, Chen, & Verkuilen, 2015, see Figure 1C).

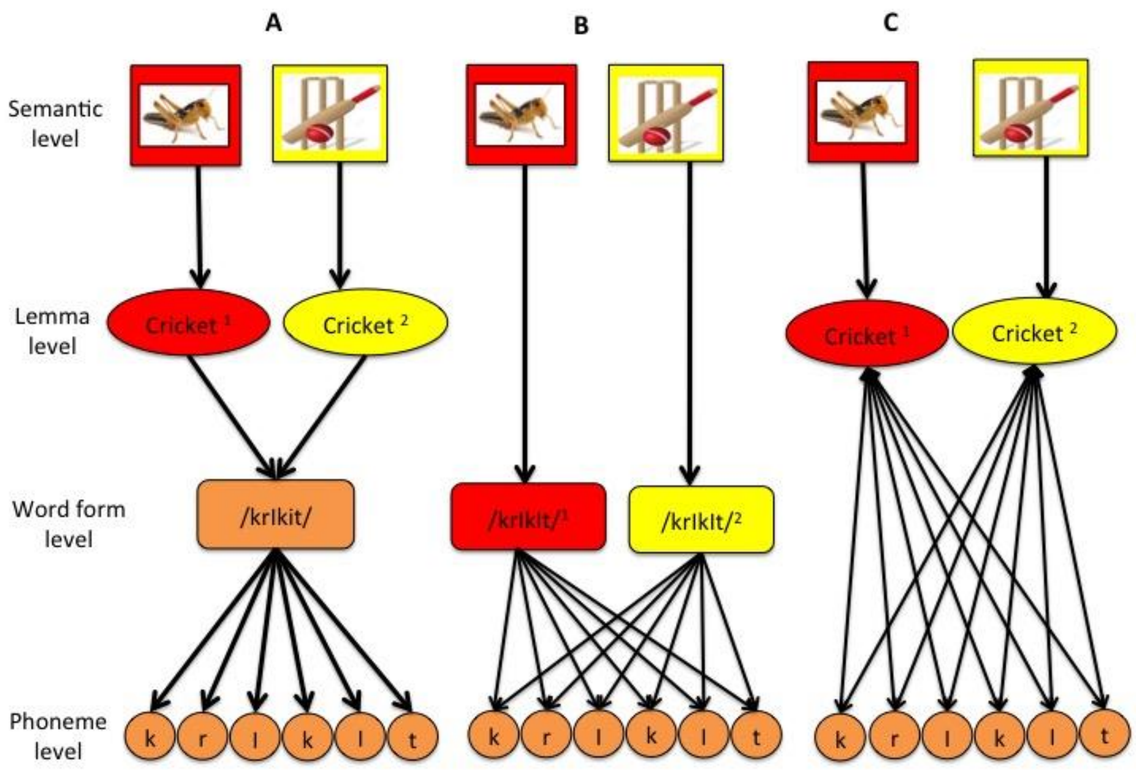


Figure 1. Homophone representation in A: Two-Stage model (Levelt et al., 1999)) B) the Independent Network model Caramazza et al.,, 2001 and C) the Interactive Activation model (Middleton et al., 2015).

Studies that have found a homophone advantage compared to non-homophones (e.g., faster translation times, Jescheniak & Levelt, 1994) have concluded that this provides evidence of a shared word form representation for homophones (e.g. Two-Stage model, Levelt et al., 1999). A shared representation will have a frequency that is the sum of both individual homophone's frequencies (frequency effects originating at the word form level). This higher frequency results in faster lexical access and production, than the frequency of each individual homophone frequency would predict (e.g., Jescheniak & Levelt, 1994).

Studies that have found no advantage for homophones (e.g., picture naming and translation, Caramazza et al., 2001; aphasic picture naming, Jacobs, Singer, & Miozzo, 2004) have attributed this to independent entries at the word form level (e.g., Caramazza et al's., 2001, Independent Network model). These studies found that homophones showed production latencies (Caramazza et al., 2001) and accuracy (Jacobs, Singer, & Miozzo, 2004) according to their individual frequencies and slower and less accurate than predicted by the summed frequency.

Due to an inconsistent landscape of evidence, the representation of homophones is still unclear. However, reviewing this previous evidence, it could be argued that the inconsistency in results might be caused by differences in methodology across studies. There are differences in the type of homophones used, the type of tasks conducted, participants' language backgrounds, methodology of comparable tasks and the control of different lexical variables in the stimuli.

Firstly, previous studies have differed in the type of homophone used. For example, Jescheniak & Levelt (1994) exclusively used homographic homophones (e.g., cricket/cricket) whereas Caramazza et al. (2001) used predominantly heterographic homophones (e.g., night/knight[3]). It is possible that differences in orthographic form of the homophones may have influenced the results, causing an advantage in some cases, and not in others (see Jescheniak and Levelt, 1994, for a homograph advantage but Caramazza et al., 2001, for no heterograph advantage). It is possible that having

[3] In the remainder of this paper, we will refer to homographic homophones as homographs and heterographic homophones as heterographs.

two separate representations in the orthographic lexicon (as heterographs must), could lead to competing activation between these representations if the orthographic lexicon was activated during phonological retrieval. This may negate any possible benefits from a shared phonological representation and result in a lack of advantage in spoken production of heterographs compared to homographs (who need not have competing representations in the orthographic lexicon). This seems plausible as many models include links between phonology and orthography (e.g. Bonin, Fayol, & Peereman, 1998; Patterson, 1986). However, Biedermann and Nickels (2008a) found no difference between heterographs and homographs in treatment generalisation in aphasia, which suggests that orthography might not be the cause of the discrepancy in these psycholinguistic studies. Nonetheless, further exploration as to the influence of orthography on the homophone advantage is required.

Another obvious difference between the previous studies is the variability of tasks used. Jescheniak and colleagues (1994; 2003) largely used translation tasks, whereas Caramazza et al. (2001) used a spoken picture naming task. This leads to the question of how far it is appropriate to directly compare the results from these different production processes. It is well established that picture naming requires prior access to semantics before the production of the phonological form (e.g., Dell et al., 1997; Levelt et al., 1991). However, it has been suggested that translation can be directly lexically mediated (Dufour & Kroll, 1995; Sholl, Sankaranarayanan, & Kroll, 1995). Indeed, evidence supports translation being less semantically mediated than picture naming (e.g., Hughes et al., 2014). This finding is interesting in the context of the inconsistency in results around the homophone effect. Indeed, Middleton et al. (2015) attributed previous inconsistent homophone production findings to homophones having a 'dual nature' depending on the stage of processing that is driving retrieval within a version of Dell and colleagues' Interactive Activation model that includes competition (Dell et al., 2004).

Middleton et al. (2015) found that in a simulation of picture naming, homophone production resulted in a frequency inheritance effect (homophones were produced more like words matched to the high frequency controls). Nevertheless, successful simulation of data from people with aphasia led them to hypothesise that during lemma retrieval (semantic-lexical retrieval - Stage 1, see Figure 1c) homophone production has a detrimental effect. Indeed, it was found that this deleterious effect of homophones was exaggerated in participants with aphasia with a deficit at this stage (compared to Stage-2 deficit participants, Middleton et al., 2015). The semantically driven nature of Stage-1 retrieval, coupled with feedback from another item with identical phonology leads to competition between homophones and interference with the selection of the correct target homophone. However, following target selection, at Stage-2 (a phonological level) reactivation of shared phonemes from the homophone partner was hypothesised to result in a homophone production advantage. Once again, this was supported by the fact that the Stage-2 impaired participants with aphasia had an exaggerated homophone advantage.

Consequently, Middleton et al. (2015) speculated that a task that has more of a semantic focus, such as naming, had a greater likelihood of no homophone advantage (despite their simulation showing an advantage), whereas a task with reduced semantic reliance and with external activation of the phonological form (such as reading aloud or translation; Jescheniak & Levelt, 1994) could result in a homophone advantage (but see Biedermann, Coltheart, Nickels, & Saunders, 2009).

Another factor contributing to the differences in results could be that some studies have involved bilingual participants and others monolinguals. Despite a wealth of bilingual research in recent years, a comprehensive production model of the bilingual lexicon is still lacking, despite the fact that being bilingual has been shown to affect both non-linguistic and linguistic abilities. For example, bilinguals have been found to have more efficient executive function abilities, reflected in, for example, better attentional control (e.g., Bialystok & Viswanathan, 2009). In contrast, for linguistic processes, monolinguals seem to show advantages in picture naming speed over bilinguals (e.g., Gollan, Montoya, Fennema-Notestine, & Morris, 2005). Therefore, it may be premature to assume the same underlying processing constraints when comparing monolingual and bilinguals speakers, even within one language. Nevertheless, Jescheniak and colleagues (1994, 2003) and Caramazza et al. (2001) *do* use bilingual (and monolingual) data in order to inform monolingual theories. Despite both studies using translation tasks there were some differences between procedures, which could have led to the differing results. As mentioned above, Jescheniak and Levelt (1994) used homographic homophones, whereas Caramazza et al. (2001; Experiment 3A) used exclusively heterographic homophones in the translation study. In addition, in Caramazza et al. (2001) participants were presented with the experimental items several times. Firstly, participants undertook a familiarisation procedure in which they were presented with the words and their translation partners to study, and then were asked to read them aloud, followed by translation of each item as fast as they could before the experiment began. The experiment was then carried out with translation occurring once more. In comparison, Jescheniak and Levelt's (1994, Experiment 6) participants only studied the items once before the experiment (but did not read aloud or practice translation); they then translated the items three times within a session averaging reaction times across these three translation times (and then deducting the reaction time on a semantic decision task for each item to give a difference score). Alario, Ferrand, Laganaro, New, Frauenfelder, and Segui, (2004) pointed out that a single familiarisation procedure reduces lexical effects, so it is possible that two prior productions of the stimuli in Caramazza et al.'s (2001) study could diminish any small homophone effects found. Indeed, Bonin, Chalard, Méot, and Fayol (2002) also suggested that pre-exposure could diminish any frequency effects.

The Present Study

The aim of this study was to examine further whether homophones show an advantage in production, and to determine the possible role of the factors, outlined above, that could have contribute to the mixed results (type of homophone, language history of participants, type of task, familiarisation). Therefore, we specifically compare monolingual and bilingual picture naming performance (Experiment 1), and compare bilingual translation tasks with bilingual picture naming (Experiment 2), contrasting heterographic and homographic words (Experiments 1 and 2), while keeping other aspects of the procedure constant (neither experiment contains a familiarisation exposure).

Experiment 1: Picture naming

Experiment 1 investigates firstly, the latency of homophone production in relation to items matched to the individual frequency of the homophone or matched to the summed frequency of the target and its homophone mate. Secondly, this experiment investigates whether monolinguals and bilinguals differ in the effects shown. Although it is known that bilinguals name pictures slightly slower than monolinguals (e.g., Gollan et al., 2005), we are interested if the *pattern* of spoken naming latencies across experimental groups differs between the participant groups. Finally, we will examine whether effects differ across homophones with identical spellings (homographs) or different spellings (heterographs). Both immediate and delayed naming data was collected. The delayed naming task was included so we could be sure that any effects found were lexical (i.e. due to frequency e.g., Savage, Bradley, & Forster, 1990). If effects are found in the immediate but not the delayed condition, and there is a significant interaction between the delayed and immediate tasks, then the influence of post-lexical artefacts can be excluded.

Method: Picture Naming

Participants

Forty-four English monolinguals and 49 Welsh-English bilinguals were recruited from Bangor University participant pool. Monolingual participants grew up in an English-only speaking home, attended an English-only school and, had lived in a monolingual community. Bilingual participants had learned both languages by the age of five and had attended primary and secondary school where teaching was delivered either bilingually or in Welsh, and were studying at Bangor University where the majority of course modules were taught in English, hence were early acquired highly proficient bilinguals. Two monolinguals were excluded due to information from the language questionnaire; they had not grown up in a monolingual English-speaking home. No bilingual participant was excluded on the basis of the questionnaires; they had all learned both languages before the age of 5 and had used both languages regularly since. Therefore the final analysis consisted of 42 monolinguals (28 females, 14 males, 19-45 years old, mean=22.32 years (SD=4.44)) and 49 bilinguals (40 females, nine males, 20-38 years old, mean age= 25.91 years (SD=5. 95)). Participants were rewarded either through course credit or seven pounds per hour.

Materials

Homophones were selected from the Alberta norms (Twilley, Dixon, Taylor, & Clark, 1994), and previous homophone studies (Biedermann & Nickels, 2008a, 2008b). Only one of each homophone pair was presented in the experiment. One of the homophone pair was chosen to be presented in the experiment if they were pictureable and lower in frequency than their partner in order to maximise any effect of frequency inheritance. The CELEX lexical database (Baayen, Piepenbrock, & Gulikers, 1996) was used to obtain written and spoken lemma and word form frequency measures. The lemma frequency measure relates to the summed frequency counts of all the derivatives of a word within that grammatical class (e.g., (a) cat, (some) cats, (the) cat's but not catty). The word-form frequency measure relates specifically to the exact form that is presented (i.e., just (a) cat). This will obviously result in different frequency counts for the same item for the lemma and word-form measures. Therefore both frequency measures were obtained and matched across groups to ensure any differences between groups were not due to matching using one frequency measure rather than the other. This resulted in 38 homophones being chosen for presentation. Eighteen of these pairs were heterographic and 20 homographic. Seventy-six non-homophonic control words were also selected from the CELEX lexical database. Non-homophonic status was determined as either having no other occurrences in the CELEX lexical database or over 95% of the frequency of all summed occurrences. Two sets of controls were created, matched pairwise to a) the (log 10 plus one) frequency of the presented homophones (individual control group) and b) the summed (log 10 plus one) frequency of both the presented and non-presented homophone partner (summed control group). Therefore, the individual control group had significantly lower frequency than the summed control group (paired sample t-tests all p>.001). Across groups the stimuli were also matched listwise on number of syllables, letters, phonemes, orthographic and phonological neighbours. These non-frequency variables for the experimental items were obtained from N-watch (Davis, 2005). See Appendix A1 for details of the characteristics of the experimental sets for the variables under consideration and B for a full list of stimuli. All stimuli were presented as photographs (300 by 300 pixels) that had at least 75% name agreement (from 15 English monolingual participants, mean age 32.5 years, who did not take part in the experiment).

Procedure

The experimental session for each participant consisted of a language questionnaire, an immediate and delayed picture naming task and a homophone definition task in this order. The questionnaires were used to assess participant's fluency and language background. The homophone definition task was administered after immediate and delayed naming tasks were completed in which participants were asked to provide definitions for both heterograph partners and one or more definitions for the homographs. If a participant could not give definitions for both homophone partners, definitions were given and participants stated if they were familiar with this meaning. If one or both meaning was not identified as familiar, this homophone pair was excluded from the analysis for that individual.

For the immediate spoken picture-naming task, participants were presented with a fixation cross in the centre of a LG 29 inch monitor PC screen for 380ms, followed immediately by a picture. The picture remained on the screen for 3000ms or until the voice key was triggered. The picture then disappeared and the next trial was initiated immediately. Participants were asked to respond as quickly and as accurately as possible. In the delayed naming task participants saw the to-be-named-picture for unlimited time. When they were ready to name, participants pressed the space bar and waited for an asterisk cue that indicated they should produce the name. The cue appeared randomly at either 1000ms or 1600ms after participants had pressed the space bar. Participants' responses and reaction times (RTs) were recorded using DMDX (Forster & Forster, 2003). As we were primarily interested in the immediate naming reaction times as a true measure of lexical access, immediate naming was always administered first. As delayed naming did not measure lexical access (or time taken to find the picture name, see above description), instead was a post-lexical access control, it was always administered after immediate naming. This is also the procedure used in other studies enlisting delayed naming as a control for lexical effects (e.g., Biedermann et al., 2009).

Results: Picture naming

Immediate and Delayed naming tasks were prepared separately. All responses were checked for correct triggers using Checkvocal (Protopapas, 2007) and scored as correct, incorrect or no response. Latencies for trials which were incorrect (8.68% immediate, 10.41% delayed), showed no response before time out (4.8% immediate, 2.10% delayed) and/ or response latencies less than 300/150ms or over 2000ms (following e.g., Barry, Hirsh, Johnston, & Williams, 2001; Starreveld, La Heij & Verdonschot 2003; 0.02% immediate, 0% delayed) were excluded from analysis (Total: immediate 13.51%, delayed 12.51%). Visual inspection of the distribution of reaction times for the 0.02% of items that were removed due to being above 2000ms confirmed that these items were outliers and not the right hand tail of the distribution. The reaction times (RTs) for which participants did not recognise both homophone meanings (0.31%) were subsequently removed. Items were removed from analysis if they were named incorrectly by over 60% of participants (in the immediate task). This resulted in removal of three homographs and one heterograph and their matched control items. For frequency, the control groups remained statistically distinct ($p<.001$ for all counts of frequency) and matched to the appropriate homophone group (all $p>.05$) and matched for all other variables (all $p>.05$; see Appendix A2 for means for the revised sets). Any RTs that were three standard deviations (SD) above or below the individual's mean were replaced with the mean (resulting in 1.39% replaced in immediate, and 1.28% replaced in delayed naming). Hence, analyses were performed on 85.09% and 86.01% of the original immediate and delayed data.

The statistical software R (R Core Team, 2015) and the *lme4* package (Bates, Mächler, Bolker, & Walker, 2014) were used to investigate the effect of language group (monolingual vs bilingual), item type (homophone, individual control, summed control), homophone spelling (heterograph vs homograph) and presentation (immediate vs delayed). Data was prepared and analysed following Baayen and Milin (2015). We inverse transformed the RT data by -0.15 as suggested by performing a Box Cox power transformation test (Box & Cox, 1964) on the immediate RTs (see Appendix C for the raw RTs). We used competitive liner mixed effects regression modelling with the immediate inverted RTs to determine the model that best fit the data in a stepwise fashion, replacing any variables that did not improve the fit of the model. This model was then used as the basis of the analysis using contrast coding to examine the nested variables of interest. The best fitting model included the four-way interaction between language group, presentation, item type and spelling, random intercepts (participants and item), name agreement, RT of the previous stimulus (one of the most significant predictors of reaction times in Baayen & Milin, 2015), visual complexity (image size) and presentation order. Number of phonemes and phonological neighbourhood did not improve the model fit further. See Appendix D for a full list of the models tested and their significance compared to the previous model. Simple effects do not inform our research questions, and are therefore not reported in detail here. However, the finding that both condition (homophone, individual or summed control) and language (monolingual, bilingual) significantly improved the model of best fit suggesting these factors affect processing. The manipulation of spelling is nested within the spelling by condition interaction[4] and is not meaningful on its own.

As the comparisons of interest were nested in our fixed factors we employed least squared means to run contrasts using the lsmeans package (Lenth, 2016). Our primary interest was in the interactions between language group, item type, spelling on the inverted RTs for immediate naming. In order to attempt to constrain the

[4] Every item (not just homophones) was assigned to a heterograph or homograph condition in the spelling variable. Hence for control items, this is meaningless (the controls for heterographs do not differ in spelling in systematic ways from the controls for homographs).

number of analyses given the problem of multiple comparisons, we restricted our analyses to those required to directly address our research questions. We ran one set of contrasts that included only immediate naming, and also a set that examined the interaction of the contrast with presentation (delayed/ immediate). A separate analysis was run on the delayed RTs. An effect would be considered to have a lexical origin if it was significant in the analysis of immediate RTs, and there was a significant interaction between that effect and delayed and immediate reaction times. Hence, here, we report results from the inverted immediate only analysis and whether this was a lexical effect. The results of the interaction with presentation (immediate x delayed) and the contrasts with the delayed presentation are reported in Appendix E. During all the contrasts we held constant the centred psycholinguistic variables that significantly predicted the inverted reaction times. We report uncorrected significance values but indicate when these did not survive correction for multiple comparisons.

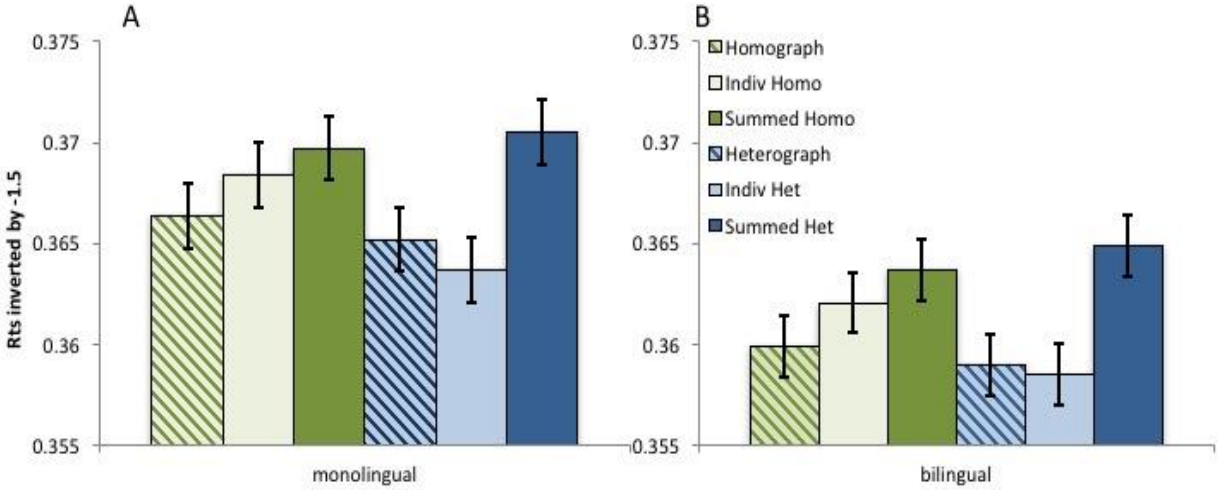


Figure 2: Response latencies (inverted RTs) for immediate picture naming. [Error bars represent standard error]

Do Homophones Show Frequency Inheritance? Analysis Of The Effect Of Condition

First we compared the two control sets to ensure that there was a significant effect of frequency on naming. As predicted, both monolinguals and bilinguals were significantly faster when naming the higher frequency summed frequency controls than lower frequency individual frequency controls ($\hat{\beta}$=-4.00 $^{e-03}$, SE = 7.85 $^{e-04}$, Z =-5.10, p<.001; $\hat{\beta}$=4.09 $^{e-03}$, SE = 8.11 $^{e-04}$, Z = 5.04, p=<.001). [5]

To examine whether homophones exhibited frequency inheritance we compared homophone naming RTs to those of the two different control sets. Both monolinguals and bilinguals named homophones slower (smaller inverted RT) than their summed frequency controls (Monolinguals: $\hat{\beta}$=4.34 $^{e-03}$, SE=8.36 $^{e-04}$, Z=5.20,

[5] To assist the reader, we repeat the statistics in the text for significant effects, but refer the reader to Table 2 for the full set of statistical results.

p<.001; Bilinguals: $\hat{\beta}$=4.81 $^{e-03}$, *SE*=8.17^{e-04}, *Z*=5.93, *p*<.001, df=193.76), as shown in Table 2 and Figure 2. In contrast, there was no significant difference between homophones and controls matched on individual frequency.

Homographs

The same pattern was found when analysing homographs separately: they were slower than their summed controls for both monolingual and bilingual immediate picture naming, although the contrast for monolinguals did not survive correction for multiple comparisons. ($\hat{\beta}$= 3.37 $^{e-03}$, *SE* = 1.15 $^{e-03}$, *Z* = 2.92, *p*<.01; $\hat{\beta}$= 3.76 $^{e-03}$, *SE* = 1. 11^{e-03}, *Z* = 3.37, *p*=<.001). As for homophones overall, monolinguals and bilinguals named homographs and their individual controls comparably.

Heterographs

The same pattern was found for heterographic homophones. Both monolinguals and bilinguals named heterographs significantly slower than controls matched to their summed frequency ($\hat{\beta}$= 5.33^{e-03}, *SE* = 1.13 $^{e-03}$, *Z* =4.70, *p*=<.001; $\hat{\beta}$= 5.90 $^{e-03}$, *SE* = 1.11 $^{e-03}$, *Z* = 5.32, *p*=<.001), and there were no significant differences in the speed of naming heterographs compared to individual controls for monolinguals or bilinguals.

Summary

This analysis finds no evidence for frequency inheritance for homophones overall, nor for heterographs or homographs separately: Homophones are named slower than their summed frequency controls but equivalent to their individual frequency controls.

Do heterographs and homographs differ in their effects? Analysis of the interaction between spelling and condition

As reported above, homographs and heterographs appeared to show similar patterns relative to their summed and individual controls. This was examined by looking first at the interaction between spelling and the comparison with summed controls, which was not significant for either monolinguals or bilinguals. While the interaction between spelling and the comparison with individual controls was closer to significance it was not significant for bilinguals and failed to survive corrections for multiple comparisons for monolinguals.

In sum, there is no strong evidence that heterographs and homographs differ in their effects in this picture naming task.

Do bilinguals and monolinguals differ in their effects? Analysis of the effect of language

As highlighted from the section above, bilinguals and monolinguals appeared to show the same effects in all comparisons. When we examined the interactions with language, none were significant: language and i) homophones vs summed controls; ii) homophone vs individual frequency; iii) homographs vs summed controls; iv) homographs vs individual controls; v) heterograph vs summed controls vi) heterograph vs individual.

Table 2. Results of the contrast analysis for picture naming

Contrast	Immediate naming	Lexical effect?
Effect of condition		
Monolingual Sum vs Indiv	$\boldsymbol{\hat{\beta}}$=4.09 $^{e-03}$, SE = 8.11 $^{e-04}$, Z = 5.04, p<.001*	
Bilingual Sum vs Indiv	$\boldsymbol{\hat{\beta}}$=-4.00 $^{e-03}$, SE = 7.85 $^{e-04}$, Z =-5.10, p<.001*	
Homophones overall (Homp= Heterographs + Homographs)		
Monolingual Homp vs Sum	$\boldsymbol{\hat{\beta}}$ =4.34 $^{e-03}$, SE=8.36 $^{e-04}$, Z=5.20, p<.001*	
Bilingual Homp vs Sum	$\boldsymbol{\hat{\beta}}$=4.81 $^{e-03}$, SE=8.17^{e-04}, Z=5.93, p<.001*	
Monolingual Homp vs Indiv	$\hat{\beta}$=-2.54 $^{e-04}$, SE = 7.96 $^{e-04}$, Z =-0.32, p=.749	n/a
Bilingual Homp vs Indiv	$\hat{\beta}$=-8.25 $^{e-04}$, SE=7. 69 $^{e-04}$, Z=-1.07, p=.285	n/a
Homographs(Homo)		
Monolingual Homo vs Sum	$\boldsymbol{\hat{\beta}}$=3.37 $^{e-03}$, SE = 1.15 $^{e-03}$, Z = 2.92, p=.004^{+}	
Bilingual Homo vs Sum	$\boldsymbol{\hat{\beta}}$=3.76 $^{e-03}$, SE = 1. 11^{e-03}, Z = 3.37, p=<.001*	
Monolingual Homo vs Indiv	$\hat{\beta}$=2.05 $^{e-03}$, SE = 1.13 $^{e-03}$, Z = 1.82, p =.071	n/a
Bilingual Homo vs Indiv	$\hat{\beta}$=2.12 $^{e-03}$, SE = 1.09 $^{e-03}$, Z = 1.95, p=.053	n/a
Heterographs (Het)		
Monolingual Het vs Sum	$\boldsymbol{\hat{\beta}}$=5.33^{e-03}, SE = 1.13 $^{e-03}$, Z =4.70, p=<.001*	
Bilingual Het vs Sum	$\boldsymbol{\hat{\beta}}$=5.90 $^{e-03}$, SE = 1.11 $^{e-03}$, Z = 5.32, p=<.001*	
Monolingual Het vs Indiv	$\hat{\beta}$=-1.54 $^{e-03}$, SE = 1.14 $^{e-03}$, Z =-1.37, p=.172	n/a
Bilingual Het vs Indiv	$\hat{\beta}$=-4.69 $^{e-04}$, SE = 1.09 $^{e-03}$, Z = -0.43, p=.667	n/a
Interaction of spelling by condition (Spelling= Het*Hom)		
Monolingual spelling* (Homp vs Sum)	$\hat{\beta}$=-9.80 $^{e-04}$, SE = 7.80 $^{e-04}$, Z =-1.26, p=.210	n/a
Bilingual spelling* (Homp vs Sum)	$\hat{\beta}$=-1.07 $^{e-03}$, SE =7.58 $^{e-04}$, Z = -1.41, p=.159	n/a
Monolingual spelling*(Homp vs Indiv)	$\hat{\beta}$=1.79 $^{e-03}$, SE =7.97 $^{e-04}$, Z=2.26, p=.025^{+}	
Bilingual spelling*(Homp vs Individual)	$\hat{\beta}$= 1.29 $^{e-03}$, SE = 7.71 $^{e-04}$, Z = 1.68, p=.095	n/a
Interaction with language (Lang= monolingual*bilingual)		
Homophones overall (Heterographs + Homographs)		
Lang x Sum vs homophone	$\hat{\beta}$=2.42 $^{e-04}$, SE=3.86 $^{e-04}$, Z=0.63, p=.530,	n/a
Lang x Indiv vs homophone	$\hat{\beta}$=2.85 $^{e-04}$, SE =3.97 $^{e-04}$, Z = 0.72, p=.472	n/a
Homographs (Homo)		
Lang x Homo vs Sum	$\hat{\beta}$= 1.96 $^{e-04}$, SE = 5.50 $^{e-04}$, Z = 0.36, p =.721	n/a
Lang x homo vs Individual	$\hat{\beta}$=3.62 $^{e-05}$, SE =5.57 $^{e-04}$, Z =0.06, p=.948	n/a
Heterographs (Het)		
Lang x Het vs Sum	$\hat{\beta}$=2.88 $^{e-04}$, SE = 5.42 $^{e-04}$, Z = 0.53, p=.595	n/a
Lang x het vs Indiv	$\hat{\beta}$= 5.35 $^{e-04}$, SE = 5.64 $^{e-04}$, Z = 0.95, p =.344	n/a

*Significance value withstood correction |+Significance did not withstand correction (Holm-bonferroni correction for 48 corrections)
Sum= control group matched to the summed frequency of the homophone
Indiv= control group matched to the individual frequency of the homophones,
Lexical effect: significant interaction between the presentation type (immediate or delayed naming)

Summary Experiment 1

In summary, both monolinguals and bilinguals named pictures of homophones, at similar speeds to control pictures matched for individual frequency of the presented homophone and significantly slower than controls matched on the combined frequency of the homophone and its higher frequency mate. In other words, we find no evidence for frequency inheritance. This effect was not modulated by homophone spelling. Although it appeared that heterographs were named slightly quicker than their individual controls and homographs slightly slower – statistically there was no evidence that heterographs and homographs showed different effects. Nor was there any effect of whether the speaker was monolingual or bilingual (although bilinguals were slower overall). This suggests that these factors are unlikely to have contributed to the inconsistent findings in the past.

Moreover, the effects that were significant in immediate production were not significant for any delayed tasks, and/or there was a significant interaction between immediate and delayed for this effect (Appendix E). Therefore, any effects found in the immediate picture naming can be attributed to lexical processing.

Experiment 2: Translation

Experiment 2 complements Experiment 1 by using a different task – cross-language translation to investigate the production of low frequency homophones compared to control items matched to either the summed frequency of the homophone and its mate or the individual frequency of that particular homophone. Once again we also investigate whether the pattern of effects is dependent on homophone spelling. The results of this experiment will also be compared to the findings from the bilingual participants in Experiment 1 in order to investigate if previous inconsistencies in the literature could be due to differences in tasks and, more specifically, if more semantically-driven tasks (such as picture naming Exp. 1) show less of a homophone advantage (or more of a disadvantage) than lexically driven tasks (such as translation Exp. 2). This could offer support for the newly proposed dual nature account.

Method: Translation

Participants

Thirty-three Welsh-English Bilingual participants, from the same population as the picture-naming participants (but not the same participants), were recruited for this study. As in the picture-naming task, participants had learnt both languages before the age of five and had continued to regularly use both languages in all aspects of life (one participant did not fulfil this criteria, so was excluded). Therefore 32 participants took part in the study in exchange for course credits (28 female, 19-51 years old, mean= 25.73 (8.31)).

Materials

Homophone pairs and control items were chosen using the same method as described in Experiment 1, however, items in this experiment also included those stimuli that were not pictureable. A Welsh speaker selected all of the items that had a different form in Welsh (i.e., items that were not loan words or cognates). This resulted in 196 Welsh items being presented for translation whose English translations were either heterographs, homographs or non-homophonic words. Unlike the picture-naming task, the experimental control groups were matched after data collection, as translation agreement for the items was not available prior to the experiment being conducted. See Appendix F for a list of the homophones, control words and translations used.

Procedure

Participants were instructed to produce the English translation of written Welsh words as fast as they could. The participants' RTs were recorded using a voice key and DMDX experimental software (Forster & Forster, 2003). Following a presentation of a fixation cross for 118ms a written Welsh word was displayed for 3000ms in the centre of the computer screen for the participants to translate orally. After the trial ended, a blank screen was displayed for 1000ms. If the voice key was triggered the word disappeared from the screen until the trial timed out. Participants were encouraged to either remain silent if they did not know the translation of the target word or to guess. The 196 items were presented in a randomised order generated by DMDX for each participant. Following the experiment, stimuli that had greater than 40% translation accuracy overall, were used to create three matched sets of homophones, summed frequency controls and individual frequency controls, with embedded subsets of heterographs and homographs. This resulted in 24 heterographs and 13 homographs each with two control sets, matched to individual and summed frequency being included in the analysis. These items did not differ significantly across subsets for any of the psycholinguistic variables presented in Appendix G.

In this experiment, unlike in Experiment 1, a delayed translation task was not included as a control for lexical effects. As Experiment 1 demonstrated that any homophone effects were truly lexical, there was no reason to believe this would be different in translation.

Results: Translation

The translation RTs were prepared and analysed in the same way as the picture naming RTs (except RTs up to 4000ms were also included in this translation experiment given the longer trial interval). All responses were corrected for mis-triggers and scored as correct, incorrect or no response using Checkvocal (Protopapas, 2007). The RTs for items where responses were incorrect, no-response, less than 300ms or over 3000ms were excluded from analysis (20.85%) in line with recent translation experiments that have used longer cut off times than picture naming (e.g., Elston- Güttler & Williams, 2008; Prior, Kroll & MacWhinney, 2013). Visual inspection and a relatively low proportion of items cut off above 3000ms (0.4%) confirmed that the RTs excluded above 3000ms reflect outliers not the right tail of the distribution. We then replaced any RTs that were three standard deviations above and below the individual mean with the mean plus/less three standard deviations (1.02%). As before, we used the *lme*4 package (Bates et al., 2014) and lsmeans package (Lenth, 2016) in R (R Core Team, 2015) to perform linear mixed regression competitive modelling and least-squared means modelling with planned contrasts.

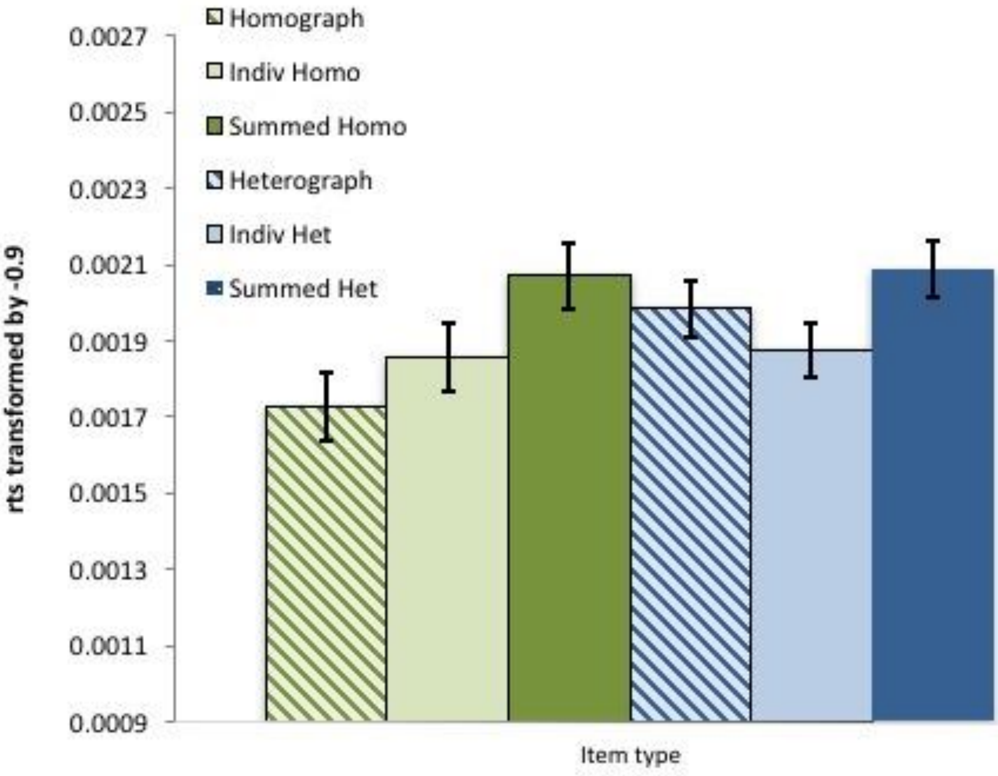


Figure 3. The translation inverted RTs transformed (-0.9), across six experimental groups with standard error bars.

The raw RTs were inverse transformed by -0.9 determined by the BoxCox transformation (Box & Cox, 1964). See Appendix H for the raw reaction times. All of the continuous variables were centred before using a step-wise competitive model procedure to determine which model to specify for the contrast analysis. The best fitting model included random slopes (items and subjects), fixed factors (the interaction between spelling and condition) and name agreement, presentation order and preceding RT. See Appendix I for all the models tested.

As noted above the simple effects are not the primary focus of this study and hence, they are not reported in detail. Nevertheless, condition (homophone, individual or summed control) significantly improved the logit mixed effect model (Appendix I). Spelling only marginally improved the model but planned contrasts are needed to fully investigate this complex nested interaction. Again, we performed planned contrasts using the lsmeans package (Lenth, 2016), holding constant the variables that influenced RTs in the logit mixed analysis. The results from the planned comparisons are presented in Figure 3 and Table 3.

Table 3. The Z-values and significance for the translation contrast analysis

Contrast	Results	Consistent with picture naming?
Effect of condition		
Homophone overall		
Homophone vs Summed	$\hat{\beta}$=-2.24 e-04 SE=6.23 e-05 Z=-3.60 p<.001	yes
Homophone vs Individual	$\hat{\beta}$=-1.07 e-05 SE= 6.32 e-05 Z= -0.17 p= .87	yes
Individual vs summed	$\hat{\beta}$=-2.14 e-04 SE=6.32 e-05 Z= -3.38 p=.001	yes
Homograph		
Homograph vs Summed	$\hat{\beta}$=-3.44 e-04 SE=1.00 e-04 Z= -3.43 p<.001	yes
Homograph vs Individual	$\hat{\beta}$=-1.30 e-04 SE= 1.02 e-04 Z= -1.28 p=.20	yes
Homographic Summed control vs Individual control	$\hat{\beta}$=-2.14 e-04 SE=1.02 e-04 Z= -2.10 p=.04	N/A
Heterograph		
Heterograph vs Summed	$\hat{\beta}$=-1.05 e-04 SE=7.38 e-05 Z= -1.42 p=.16	no
Heterograph vs Individual	$\hat{\beta}$ =1.09 e-04 SE=7.43 e-05 Z= 1.46 p= .15	yes
Heterographic Summed control vs Individual control	$\hat{\beta}$=-2.13 e-04 SE=7.44 e-05 Z= -2.87 p= .005	N/A
Spelling by condition		
Spelling*(homophone vs summed)med control	$\hat{\beta}$=1.20 e-04 SE=6.23 e-05 Z= 1.92 p=.06	yes
Spelling* (homophone vs individual)	$\hat{\beta}$=1.20 e-04 SE=6.29 e-05 Z= 1.90 p=.06	yes
Spelling*(homophone vs summed) *(homophone vs individual control)	$\hat{\beta}$=2.19 e-07 SE=6.30 e-05 Z= 0.00 p= 1.00	N/A

Homophones= heterographs and homographs combined
Individual= controls matched to the individual homophone frequency
Summed= controls matched to the sum of homophone frequency
Spelling= Heterographic or homographic homophones

Do Homophones Show Frequency Inheritance? Analysis of the Effect of Condition

Homophones overall were translated significantly slower than controls matched to the summed frequency of both homophone meanings ($\hat{\beta}$=-2.24 e-04, SE=6.23 e-05, Z=-3.60, p<.001), but there was no significant difference between RTs for homophones and the controls matched to their individual frequency.

Homographs showed the same pattern: they were translated significantly slower than their summed controls ($\hat{\beta}$=-3.44 $^{e-04}$, *SE*=1.00 $^{e-04}$, *Z*= -3.429, *p*<.001) but not significantly different from their individual frequency matched controls.

In contrast, for heterographs there was no difference in translation speed for heterographs and either their summed controls, or their individual frequency matched controls.

Do Heterographs and Homographs Differ in their Effects? Analysis of the interaction between Spelling and Condition

The interactions between the spelling of homophones and their RT compared with both controls dependent on spelling were both approaching significance (summed:$\hat{\beta}$=1.20 $^{e-04}$, *SE*=6.23 $^{e-05}$, *Z*= 1.92, *p*=.06; individual: $\hat{\beta}$=1.20 $^{e-04}$, *SE*=6.29 $^{e-05}$, *Z*= 1.90, *p*=.06): Compared to summed controls there was a tendency for homographs to be named slower than heterographs. In contrast, compared to their individual controls, homographs tended to be named slower while heterographs tended to be named faster. The interaction between homophones dependent on spelling and type of control (summed or individual) was not significant.

Summary

In this translation task, we once again have no evidence for frequency inheritance for homophones overall: Homophones have translation times which are similar to individual frequency matched controls and significantly slower than controls matched on the summed frequency in. Once again, there were no significant differences in the patterns depending on the spelling of the homophone, however, the interactions were close to significant reflecting that while homographs were slightly (but not significantly) slower than their individual frequency controls, heterographs had reaction times intermediate between the summed and individual frequency matched controls.

General Discussion

The aim of this study was to determine the possible role of the factors, outlined above, that could have contributed to the mixed results of previous studies (type of homophone, language history of participants, type of task, familiarisation). The study investigated the nature of homophone representation in the phonological output lexicon, by studying whether homophones show an advantage in production compared to non-homophones. We focused on three different factors that we hypothesised may have contributed to the mixed results reported in the literature: i) whether speakers are monolingual or bilingual, ii) whether homophones are homographic and share a spelling (e.g., *calf*- infant cow, *calf*-leg muscle) or heterographic with distinct spellings (e.g., *doe*- deer, *dough*-bread and iii) whether the task involves greater semantic mediation (as in picture naming), or less semantic mediation (as in translation).

Overall, we found a consistent pattern that homophones were produced no differently in speed to control words of equivalent frequency (individual frequency matched control). In contrast, in general, they were produced significantly slower than control items of higher frequency (that were matched to the summed frequency of both homophone meanings). This was observed for both speaker groups (mono- and bilingual), and this was evident for both immediate picture naming and translation (but not in delayed naming confirming the lexical origin of the differences between conditions).

Hence, at this level of analysis, we have no evidence that homophones show 'frequency inheritance' – they are not responded to faster than is predicted by their individual frequency. Nor are they produced as fast as the sum of the frequency of both homophones as would be predicted if homophones shared a representation. This supports views where the word form representation is not shared, such as Caramazza et al.'s (2001) Independent Network model. It can also be consistent with Middleton et al.'s (2015) account where there are both independent representations (which compete) and feedback (which facilitates) resulting in no net gain for homophone production. We will return to this below, but first examine the role of the factors manipulated within and across experiments.

(i) Does bilingualism affect the processing of homophones?

Overall bilinguals were slower naming pictures compared to monolinguals. This is in line with previous results and is thought to be a lexical effect (e.g., Gollan et al., 2005). However despite an overall effect of bilingualism, bilinguals and monolinguals showed the same patterns when naming homophones and their controls. Our study finds no evidence to support a position where differences between experiments on homophone naming could be due to the language status of the participant. Nevertheless, our results do not exclude the possibility that bilinguals who acquire their second language later in life and/or are less proficient could show different effects. This is a topic that future research could address.

(ii) Does homophone spelling affect homophone production?

Regardless of spelling, both types of homophone were produced with no significant difference in speed compared to their individual frequency controls in both picture naming and translation. Interestingly, heterographs showed a (non-significant) tendency to be produced faster and homographs produced slower than their individual controls, across both tasks and participant groups.

In relation to summed controls, although homophones combined were produced significantly slower than summed controls, the pattern was less clear when heterographs and homographs were analysed separately: heterographs were significantly slower than summed controls for both monolinguals and bilinguals in naming, but not significantly slower for translation (although the pattern was in the same direction). In contrast, homographs were significantly different in the translation task, and in naming for bilinguals, but the effect did not withstand correction for multiple comparisons for monolingual picture naming. In fact, for the translation task, the interaction between spelling (homograph vs heterograph) and the contrast with summed controls was approaching significance, as was the interaction between spelling and individual controls.

This provides some, albeit weak, suggestion there could be differences between heterographic and homographic homophone processing. As noted in the Introduction, Jescheniak and Levelt (1994) and Caramazza et al. (2001) Jescheniak and Levelt (1994) and Caramazza et al. (2001) showed differing translation effects of homophones, and it has been suggested that the type of homophone may have played a role in this. Jescheniak and Levelt (1994) detected no difference between translation times for summed controls and *homographs* and both of these were produced faster than the individual controls. Whereas Caramazza et al. (2001) discovered significant differences between summed controls and *heterographs,* which were not significantly different from individual controls. In contrast, however, we found almost the opposite: homographs were produced more similarly to individual controls and these were significantly slower than the summed controls, and did not find any difference between heterographs and summed controls. Consequently, while it is clear that there is inconsistency, it cannot be attributed in any systematic way to differences between heterographic and homographic stimuli. Perhaps the inconsistency could be due to the differences in homophones in the different languages used, or even differences in translating or psycholinguistic variables between languages.

(iii) Does the context in which homophone words are produced (picture naming vs translation) affect their production?

The overall patterns were very similar for the translation of homophones and controls and the naming of pictures: Homophones are not significantly different in production latencies to their individual controls but are slower than their summed controls. Likewise, comparing Figures 2 and 3 (earlier) the patterns shown by heterographs and homographs are very similar across the two tasks. However, there may be some indication that the effects differed in strength across tasks. For example, there was only a significant difference between the heterographs and their summed controls in the picture naming but in the translation task, there was no significant difference between the heterographs and either control group.

How might these results relate to Middleton et al.'s (2015) dual nature account of homophony? Middleton et al. (2015) postulated that whether or not there is a homophone benefit in production depends on the balance between competition at the first stage of retrieval and facilitation in the second stage of retrieval. Middleton et al. argued that in semantically driven tasks (like picture naming) there would be a greater influence of competition between the homophone nodes as a result of feedback from the shared segment (phoneme) nodes. Hence, this predicts that we should see homophones showing relatively more of a disadvantage, or less of an advantage, relative to matched controls in the naming task than the translation task. In some aspects our data, could be compatible with Middleton et al.'s account. Although not statistically compared across tasks, heterographs do indeed seem to show a relative benefit in the translation task – they were significantly slower than summed controls in naming, but not significantly slower for translation (although the pattern was in the same direction). However, the homographs showed the same pattern across both tasks. Is it possible that it is only when there relatively less competition due to the inconsistent spelling for heterographs, that the benefit of reduced semantic processing in translation is apparent. This relies on the phonological lexicon being influenced by the orthographic lexicon. Activation from one orthographic homographic representation (parallel activation of orthography during phonological production has been debated see Bonin, Fayol, & Gombert, 1997) could spread to two representations in the phonological lexicon. Whereas, activation of one heterograph would spread to a single phonological representation, resulting in either facilitation or less competition. Regardless of how, this might explain the slight advantage heterographs had over homographs, as translation is a more lexically mediated task. Alternatively, it may be that translation includes sufficient semantic activation to prevent the difference between tasks. However, simulation of whether there is a difference in the predictions made by the Dual Nature account depending on the spelling of homophones would be worth further investigation.

Of course our study is not without limitations, one methodological flaw of the translation task needs addressing in the future: sets were not matched for psycholinguistic variables across languages in the translation task. This was both because we were replicating Jescheniak and Levelt (1994), who did not cross-language match, but also because very few lexical databases are comparable across languages. However it seems plausible, as they were matched in English, the translation sets may have been matched on psycholinguistic variables in Welsh. Nevertheless, studies investigating reading (within a language) find frequency effects (e.g., Gerhand & Barry, 1998) which suggest the time taken to read the word before translation could be affecting reaction times. A translation study that was able to match cross-linguistic variables within and across languages could more confidently interpret any homophone effects. Nevertheless, our naming and translation experiments were consistent in their findings suggesting that similar factors were at play. In addition, although many studies have not controlled whether homophone partners share a word class or not, there is some evidence suggesting differing effects of homophones depending on word class (e.g., White, Abrams, McWhite, & Hagler, 2010). Perhaps the results of the present study would be different if only within word class (e.g. noun-noun) homophones had been used. Indeed, we considered this possibility, but the limited number of pictureable homophones available meant that this simply was not possible.

Another possibility is that differences in the size of the difference between the summed and individual frequencies of homophones may have been insufficient. Indeed, looking at the log frequency difference, while this was significant, there seems to be less of a difference between the summed homophone and control frequencies and the individual homophone frequencies in the present study compared to previous frequency inheritance studies (e.g., Caramazza et al., 2001; Jescheniak & Levelt, 1994). This could explain the reduced effects currently presented; perhaps the frequency inherited needs to be considerable to measure any reaction time benefit. However, Caramazza et al. (2001) had larger frequency differences than the present study and yet found no significant frequency inheritance. Hence it seems that lack of frequency inheritance cannot always be attributed to small frequency differences.

Conclusions

The present study attempted to investigate whether homophones share a word form representation, have independent word form representations, or separate lemma representations in an interactive model as in the models specified in Figure 1. We found that variation in previous homophone research was unlikely to be due to differences in participants (monolingual or bilingual). Although the overall pattern of naming heterographs, homographs and their controls remains the same across tasks, whether or not the difference between heterographs and summed controls was significant depends on the task. Therefore, future unimpaired studies should control for homophone spelling. However, our data also suggest that results of studies examining translation of heterographs and homographs are difficult to replicate. Therefore, we need to take caution when drawing conclusions from spoken homophone production studies in unimpaired speakers.

In summary, our findings add to the previously found inconsistences. It seems that, depending on the conditions of the experiments, homophone production results in an advantage or disadvantage (compared to summed or individual frequency matched controls). Middleton et al's. (2015) Dual Nature account is the first theory that can explain both advantageous and deleterious results of homophone production. This appears to be the theory with the most promise in potentially accounting for the present and previously found results.

References

Alario, F.-X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers*, *36*(1), 140–155. https://doi.org/10.3758/BF03195559

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1996). Celex2. *Linguistic Data Consortium, Philadelphia*.

Barry, C., Hirsh, K. W., Johnston, R. A., & Williams, C. L. (2001). Age of Acquisition, Word Frequency, and the Locus of Repetition Priming of Picture Naming. *Journal of Memory and Language*, *44*(3), 350–375. https://doi.org/10.1006/jmla.2000.2743

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting Linear Mixed-Effects Models using lme4. *arXiv:1406.5823 [Stat]*. Retrieved from http://arxiv.org/abs/1406.5823

Bialystok, E., & Viswanathan, M. (2009). Components of executive control with advantages for bilingual children in two cultures. *Cognition*, *112*(3), 494–500. https://doi.org/10.1016/j.cognition.2009.06.014

Biedermann, B., Coltheart, M., Nickels, L., & Saunders, S. (2009). Effects of homophony on reading aloud: Implications for models of speech production. *Language and Cognitive Processes*, *24*(6), 804–842. https://doi.org/10.1080/01690960802597243

Biedermann, B., & Nickels, L. (2008a). Homographic and heterographic homophones in speech production: Does orthography matter? *Cortex*, *44*(6), 683–697. https://doi.org/10.1016/j.cortex.2006.12.001

Biedermann, B., & Nickels, L. (2008b). The representation of homophones: More evidence from the remediation of anomia. *Cortex*, *44*(3), 276–293. https://doi.org/10.1016/j.cortex.2006.07.004

Bonin, P., Chalard, M., Méot, A., & Fayol, M. (2002). The determinants of spoken and written picture naming latencies. *British Journal of Psychology*, *93*(1), 89–114. https://doi.org/10.1348/000712602162463

Bonin, P., Fayol, M., & Gombert, J.-E. (1997). Role of phonological and orthographic codes in picture naming and writing: An interference paradigm study. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, *16*(3), 299–324.

Bonin, P., Fayol, M., & Peereman, R. (1998). Masked form priming in writing words from pictures: Evidence for direct retrieval of orthographic codes. *Acta Psychologica*, *99*(3), 311–328. https://doi.org/10.1016/S0001-6918(98)00017-1

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological*, 211–252.

Caramazza, A., Costa, A., Miozzo, M., & Bi, Y. (2001). The specific-word frequency effect: Implications for the representation of homophones in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(6), 1430–1450. https://doi.org/10.1037/0278-7393.27.6.1430

Davis, C. J. (2005). N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods*, *37*(1), 65–70. https://doi.org/10.3758/BF03206399

Dell, G. S., Lawler, E. N., Harris, H. D., & Gordon, J. K. (2004). Models of errors of omission in aphasic naming. *Cognitive Neuropsychology*, *21*(2–4), 125–145. https://doi.org/10.1080/02643290342000320

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, *104*(4), 801–838. https://doi.org/10.1037/0033-295X.104.4.801

Dufour, R., & Kroll, J. F. (1995). Matching words to concepts in two languages: A test of the concept mediation model of bilingual representation. *Memory & Cognition*, *23*(2), 166–180. https://doi.org/10.3758/BF03197219

Elston-Güttler, K.E., Williams, J.N., (2008) First language polysemy affects second language meaning interpretation: evidence for activation of first language concepts during second language reading*. Second Language Research 24*(2), 167–187

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*(1), 116–124. https://doi.org/10.3758/BF03195503

Gerhand, S., & Barry, C. (1998). Word frequency effects in oral reading are not merely age-of-acquisition effects in disguise.

Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition*, *33*(7), 1220-1234.

Hughes, E. K., Roberts, R. J., Roberts, D. J., Kendrick, L. T., Payne, J. S., Owen-Booth, B., … Tainturier, M.-J. (2014). Is translation semantically mediated? Evidence from Welsh-English bilingual aphasia. *Frontiers in Psychology: Academy of Aphasia- 52nd Annual Meeting*. https://doi.org/10.3389/conf.fpsyg.2014.64.00082

Jacobs, M., Singer, N., & Miozzo, M. (2004). The representation of homophones: Evidence from anomia. *Cognitive Neuropsychology*, *21*(8), 840–866. https://doi.org/10.1080/02643290342000573

Jaeger, T. F. (2008). Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of Memory and Language*, *59*(4), 434–446. https://doi.org/10.1016/j.jml.2007.11.007

Jescheniak, J. D., & Levelt, M., J. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 824–843. https://doi.org/10.1037/0278-7393.20.4.824

La Heij, W., Heikoop, K.W., Akerboom, S., Bloem, I, (2003). Picture naming in picture context: Semantic interference or semantic facilitation? *Psychology* Science, 45, (1),. 49-62

Lenth, R. (2016). lsmeans: Least-Squares Means (Version 2.23-5). Retrieved from https://cran.r-project.org/web/packages/lsmeans/index.html

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(01), 1–38.

Levelt, W. J. M., Schriefers, H., Meyer, A. S., Pechmann, T., Vorberg, D., Braunschweig, T. U. it, & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review*, *98*, 122–142.

Middleton, E. L., Chen, Q., & Verkuilen, J. (2015). Friends and foes in the lexicon: Homophone naming in aphasia. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(1), 77–94. https://doi.org/10.1037/a0037778

Patterson, K. (1986). Lexical but nonsemantic spelling? *Cognitive Neuropsychology*, *3*(3), 341–367. https://doi.org/10.1080/02643298608253363

Prior, A., Kroll, J.F., MacWhinney, B., (2013). Translation ambiguity but not word class predicts translation performance. *Bilingualism: Language and Cognition 16* (2), 2013, 458–474 C Cambridge University Press 2012 doi:10.1017/S1366728912000272

Protopapas, A. (2007). Check Vocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods*, *39*(4), 859–862. https://doi.org/10.3758/BF03192979

R Core Team. (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/. Retrieved 17 August 2016, from https://www.google.com.au/search?q=R+Core+Team+(2015).+R%3A+A+language+and+environment+for+statistical+computing.+R+Foundation+for+Statistical+Computing%2C+Vienna%2C+Austria.+URL+https%3A%2F%2Fwww.R-project.org%2F.&ie=utf-8&oe=utf-8&client=firefox-b-ab&gfe_rd=cr&ei=I_C0V6uIGMLr8AeMwrvABw

Savage, G. R., Bradley, D. C., & Forster, K. I. (1990). Word frequency and the pronunciation task: The contribution of articulatory fluency. *Language and Cognitive Processes*, *5*(3), 203–236. https://doi.org/10.1080/01690969008402105

Schwartz, M. F., Dell, G. S., Martin, N., Gahl, S., & Sobel, P. (2006). A case-series test of the interactive two-step model of lexical access: Evidence from picture naming. *Journal of Memory and Language*, *54*(2), 228–264. https://doi.org/10.1016/j.jml.2005.10.001

Sholl, A., Sankaranarayanan, A., & Kroll, J. F. (1995). Transfer Between Picture Naming and Translation: A Test of Asymmetries in Bilingual Memory. *Psychological Science*, *6*(1), 45–49. https://doi.org/10.1111/j.1467-9280.1995.tb00303.x

Twilley, L. C., Dixon, P., Taylor, D., & Clark, K. (1994). University of Alberta norms of relative meaning frequency for 566 homographs. *Memory & Cognition*, *22*(1), 111–126. https://doi.org/10.3758/BF03202766

Wheeldon, L. R., & Monsell, S. (1992). The Locus of Repetition Priming of Spoken Word Production. *The Quarterly Journal of Experimental Psychology Section A*, *44*(4), 723–761. https://doi.org/10.1080/14640749208401307

White, K. K., Abrams, L., McWhite, C. B., & Hagler, H. L. (2010). Syntactic constraints in the retrieval of homophone orthography. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *36*(1), 160–169. https://doi.org/10.1037/a0017676

Appendices

Table A1. Mean frequency (log10) and other psycholinguistic variables by experimental groups for the picture naming task before data collection.

		All Homophones & Controls				Homographs (hom) & Controls				Heterograph(het) & Controls			
		N=38				N=20				N=18			
		Indiv	Indiv Control	Sum	Sum control	Indiv homo	Indiv Control	Sum homo	Sum control	Indiv het	Indiv control	Sum het	Sum control
Frequency	Spoken word form	0.89*	0.92*	1.41+	1.38+	0.79*	0.87	1.17+	1.33+	0.99*	0.97*	1.66+	1.37+
	Written word form	2.25*	2.55*	2.85+	2.96+	2.02*	2.54	2.57+	2.83+	2.51*	2.57*	3.16+	2.94+
	Spoken lemma	1.01*	0.94*	1.63+	1.46+	0.91*	0.86*	1.47+	1.55+	1.13*	1.03*	1.81+	1.45+
	Written Lemma	2.54*	2.61*	3.10+	3.06+	2.43*	2.57*	2.89+	3.02+	2.67*	2.57*	3.33+	3.08+
Syllables		1.13	1.11	1.13	1.21	1.18	1.1	1.18	1.3	1.11	1.12	1.11	1.17
Phonemes		3.24	3.24	3.21	3.06	3.41	3.25	3.41	3.6	3.11	3.22	3.11	3.17
Letters		4.26	4.39	4.18	4.29	4.24	4.45	4.24	4.45	4.33	4.33	4.17	4.22
Orth N		7.63	5.71*	8.97	8.13	8.88	5.5	8.88	7	5.78	5.94	8.61	9.61+
Phon N		18.37	15.45*	18.45	17.92	17.53	14.15	17.53	16.4	18.11	16.89	18.11	18.83
Phon N frequency		306.47	239.24	335.1	236.6	255.63	114.61	255.63	117.66	668.95	373.12	668.95	168.68+

*significantly different from the summed homophone group (or summed heterograph/ summed homograph frequency) in a 1-tailed paired t-test (p<.05)

+ significantly different from the individual homophone group(or summed heterograph/ summed homograph frequency) in a 1-tailed paired t-test(p<.05)

Indiv= Individual
Orth N= Number of orthographic neighbours
Phon N= Number of phonological neighbours
Phon N frequency= average frequency of phonological neighbours

Table A2. Mean frequency (log10) and other psycholinguistic variables by experimental groups for the picture naming task after data collection.

		All Homophones & Controls N=34				Homographs (homo) & Controls N=17				Heterograph (het) & Controls N=17			
		Indiv	Indiv Control	Sum	Sum control	Indiv homo	Indiv control	Sum homo	Sum control	Indiv het	Indiv control	Sum het	Sum control
Frequency	Spoken word form	0.94*	0.84*	1.49+	1.46+	0.87*	0.83*	1.30+	1.45+	1.02*	0.85*	1.68+	1.48+
	Written word form	2.36*	2.47*	2.97+	2.95+	2.23*	2.52	2.77+	2.89+	2.50*	2.42*	3.18+	3.01
	Spoken lemma	1.05*	0.90*	1.65+	1.57+	0.94*	0.89*	1.48+	1.55+	1.16*	0.92*	1.82+	1.59+
	Written Lemma	2.54*	2.59*	3.13+	3.07+	2.41*	2.62*	2.93+	3.02+	2.67*	2.57*	3.33+	3.12+
Syllables		1.12	1.12	1.12	1.24	1.18	1.12	1.18	1.29	1.06	1.12	1.06	1.18
Phonemes		3.18	3.29	3.18	3.38	3.41	3.24	3.41	3.59	2.94	3.35	2.94	3.18
Letters		4.26	4.35	4.15	4.29	4.24	4.47	4.24	4.35	4.29	4.24	4.06	4.42
Orth N		7.35	6	8.79	8.53	8.71	5.47	8.71	7.88	6	6.53	8.88	9.18
Phon N		17.97	15.24	17.97	18.18	16.76	13.35	16.76	17.18	19.18	17.12	19.18	19.18

*significantly different from the summed homophone group (or summed heterograph/ summed homograph frequency) in a 1-tailed paired t-test (p<.05)

+significantly different from the individual homophone group(or summed heterograph/ summed homograph frequency) in a 1-tailed paired t-test(p<.05)

Indiv= Individual

Sum= the sum of both homophone partners frequency

Orth N= Number of orthographic neighbours

Phon N= Number of phonological neighbours

Appendix B. A list of all the stimuli for Experiment 1

Presented homophone	Individual frequency control	Summed frequency control
Homograph (and controls)		
Calf (muscle)	Owl	Rope
Can (tin)	Broom	Sock
Bark (tree)	Barn	Goat
Bow (ribbon)	Shark	Skull
Mole (face)	Ghost	Fox
Lace (fabric)	Cork	Shoulder
Chest (body)	Apple	Plate
Ruler (measure)	Lamp	Beard
Scale (fish)	Tent	Ice
Leaf (tree)	Sleeve	Roof
Fly (bug)	Tongue	Hat
Band (music)	Shirt	Bell
Cross (crucifix)	Arrow	Foot
Iron (clothes)	Sheep	Paper
Watch (time)	Path	Radio
Bridge (water)	Van	Car
Fire (flames)	Bus	Money
Card* (paper)	Knee*	Church*
Duck* (under)	King*	Jacket*
Well* (water)	Sword*	Teeth*
Heterographs (and controls)		
Buoy	Raft	Bone
Dough	Moth	Scarf
Waist	Peach	Lamb
Pear	Wig	Cheese
Nun	Stool	Cat
Veil	Rug	Cow
Shoe	Kite	Clock
Fur	Wool	Wing
Ceiling	Bench	Ear
Bee	Monkey	Bag
Beach	Collar	Nose
Knight	Leg	Horse
Key	Wheat	City
Plane	Bird	Table
Floor	Rice	Coffee
Hair	Arm	Door
Sun	Boat	Book
Flour*	Vase*	Pill*

* removed before analysis due to less than 40% accuracy and/or matching purposes

Appendix C. The mean raw reaction times (standard deviations) for the Picture naming study

	Monolingual		Bilingual	
	Homograph	Heterograph	Homograph	Heterograph
Homophone	848 (213)	866(243)	954 (256)	970 (275)
Individual control	805 (215)	897(237)	908 (249)	986 (256)
Summed control	778(192)	757 (183)	873 (230)	844 (222)

Appendix D. The various models tested and the p values and chi squared (Chisq) values from ANOVAs comparing each model with the previously significant one using the picture naming inverted RTS

	Model	Chisq	sig
1	(1|subject)+(1|items)		
2	(condition)+(1|subject)+(1|items)	23.65	.001
3	(Language+condition)+(1|subject)+(1|items)	5.96	0.01
4	(Spelling+Language +condition)+(1|subject)+(1|items)	21.37	<.001
5	(Spelling+Language+condition+presentation)+(1|subject)+(1|items)	18175	0.36
6	(Spelling+Language+condition+presentation)^4+(1|subject)+(1|ite ms)	208.02	<.001
7	(Spelling+Language+condition+presentation)^4+ NA +(1|subject)+(1|items)	48.55	<.001
8	(Spelling+Language+condition+presentation)^4+ NA +Phoneme+(1|subject)+(1|items)	0.16	0.69
9	(Spelling+Language+condition+presentation)^4+ NA +Phon N+(1|subject)+(1|items)	0.03	0.85
10	(Spelling+Language+condition+presentation)^4+ NA +preceedingRT+(1|subject)+(1|items)	23.1	<.001
11	(Spelling+Language+condition+presentation)^4+ NA +preceedingRT+Original+(1|subject)+(1|items)	6.55	0.01
12	((Spelling+Language+condition+presentation)^4+NA+preceedingR T+Original+Order+(1|subject)+(1|items)	12.81	<.001

Spelling= heteterograph, homograph
Language- monolingual, bilingual

Condition= homophone, individual frequency matched control, summed frequency matched control
NA= Nameagreement i.e., percentage of correct name produced (centred)

Phoneme= number of phonemes (centred)
Phon N= number of phonological neighbours (centred)
preceedingRT= the latency of the preceding item (centred)
Original= size of the image (a visual complexity measure) (centred)
Order= the presentation order (centred)

Appendix E. The results of contrasts for the delayed and immediate by delayed interaction (Imm*delayed)

Contrast	Delayed naming	Imm*deleayed	Lexical effect?
Effect of condition			
Monolingual Sum vs Individual	β =-1.42 $^{e-04}$, *SE* = 8.61 $^{e-04}$, *Z* =-0.16, *p*=.87	β = -3.16 $^{e-03}$, *SE* = 7.73 $^{e-04}$, *Z* = -4.09, *p* <.001*	
Bilingual Sum vs Individual	β =3.72 $^{e-04}$, *SE* = 8.43 $^{e-04}$, *Z* =0.44, *p*=.66	β = 3.46 $^{e-03}$, *SE* = 7.44 $^{e-04}$, *Z* = 4.65, *p*<.001*	
Homophones overall (Homophones= homographs + heterographs)			
Monolingual Homophones vs Sum	β =1.72 $^{e-03}$, *SE*=7.77 $^{e-04}$, *Z*=2.21, *p*=.027+	β =2.71 $^{e-03}$, *SE*=7.92 $^{e-04}$, *z*=-3.48, *p*=.002+	
Bilingual Homophone vs Sum	β =1.44 $^{e-03}$, *SE*=7.60 $^{e-04}$, *Z*=1.90, *p*=.058	β =-4.28 $^{e-03}$, *SE*=7.45 $^{e-04}$, *Z*=-5.78, *p*<.001*	
Monolingual Homophones vs Individual	β =-1.86 $^{e-03}$, *SE* = 7.40 $^{e-04}$, *Z* =-2.51, *p*=.012	β =-2.22 $^{e-04}$, *SE*=7.88 $^{e-04}$, *Z*=-0.28, *p*=.78	n/a
Bilingual Homophone vs Individual	β =-1.81 $^{e-03}$, *SE*=7. 18 $^{e-04}$, *Z*=-2.52, *p*=.012+	β = 5.41 $^{e-04}$, *SE* = 7.58 $^{e-04}$, *Z* =0.71, *p* = .475	n/a
Homographs (homo)			
Monolingual Homo vs Sum	β =9.46 $^{e-04}$, *SE* = 1.01 $^{e-03}$, *Z* = 0.93, *p*=.35	β =4.07 $^{e-03}$, *SE* = 1.15 $^{e-03}$, *Z* = 3.54, *p*=<.001*	
Bilingual Homo vs Sum	β =1.27 $^{e-03}$, *SE* = 9. 78 $^{e-03}$, *Z* = 1.30, *p*=.19	β = 2.78 $^{e-03}$, *SE* = 1.08 $^{e-03}$, *Z* = 2.58, *p* = .010 $^{+}$	n/a
Monolingual Homo vs Individual	β =1.90 $^{e-03}$, *SE* = 1.12 $^{e-03}$, *Z* = 1.7, *p* =.09	β = -2.48 $^{e-03}$, *SE* = 1.15 $^{e-03}$, *Z* = -2.14, *p* =.03+	n/a
Bilingual Homo vs Individual	β =1.65 $^{e-03}$, *SE* = 1.07 $^{e-03}$, *Z* = 1.53, *p*=.13	β = -3.86 $^{e-03}$, *SE* = 1.07 $^{e-03}$, *Z* =-3.62, *p*=.001*	n/a
Heterographs (het)			
Monolingual Het vs Sum	β =2.48 $^{e-03}$, *SE* = 1.11 $^{e-03}$, *Z* =2.24, *p*=.025+	β = 8.97 $^{e-03}$, *SE* = 1.97 $^{e-04}$, *Z* = 4.56, p<.001*	
Bilingual Het vs Sum	β =1.62 $^{e-03}$, *SE* = 1.09 $^{e-03}$, *Z* = 1.47, *p*=.140	β = 3.69 $^{e-03}$, *SE* = 1.09 $^{e-03}$, *Z* = -3.39, *p* <.001*	
Monolingual Het vs Individual	=1.82 $^{e-03}$, *SE* = 1.03 $^{e-03}$, *Z* =1.76, *p*=.078,	β =-2.66 $^{e-03}$, *SE* =1.35 $^{e-03}$, *Z* =-2.34 , *p*= .02 $^{+}$,	n/a
Bilingual Het vs Individual	β =1.98 $^{e-03}$, *SE* = 1.00 $^{e-03}$, *Z* = 1.97, *p*=.048+	β =-1.86 $^{e-03}$, *SE* = 5.28 $^{e-04}$ *Z* =-3.53, *p*=<.001*	n/a
Interaction of spelling by condition (Spelling= homograph* heterography)			
Monolingual spelling* (Homophone vs Sum)	β =-7.69 $^{e-04}$, *SE* =7.24 $^{e-04}$, *Z*=-1.06, *p*=.29	β =-2.00 $^{e-04}$, *SE* = 7.93 $^{e-04}$, *Z* =-0.25, *p*=.80	n/a
Bilingual spelling* (Homophone vs Sum)	β = -1.73 $^{e-04}$, *SE* = 7.07 $^{e-04}$, *Z* = -0.24, *p*=.81	β = 2.67 $^{e-04}$, *SE* =7.46, *Z* = 0.36, *p*=0.720	n/a
Monolingual spelling*(Homophone vs Individual)	β =3.89 $^{e-05}$, *SE* =7.80 $^{e-04}$, *Z*=0.05, *p*=.96	β =-3.27 $^{e-03}$, *SE*= 8.16 $^{e-04}$, *Z*=-4.01, *p*<.001*	**
Bilingual spelling* (Homophone vs Individual)	β = 1.68 $^{e-04}$, *SE* = 7.52 $^{e-04}$, *Z* = -0.22, *p*=.82	β =-3.10 $^{e-03}$, *SE*=1.20 $^{e-03}$, *Z*=-2.58, *p*=.001 $^{+}$	n/a
Interaction with language (lang= monolingual *bilingual)			
Homophones overall (Homophones= homographs + heterographs)			
Lang x Sum vs homophone	β =-0.37 $^{e-04}$, *SE*=3.92 $^{e-04}$, *Z*=-0.35, *p*=.73	β = 7.69 $^{e-04}$, *SE*=5.43 $^{e-04}$, *Z*=-1.42, *p*=.16	n/a
Lang x Individual vs homophone	β =2.21 $^{e-05}$, *SE* =4.07 $^{e-04}$, *Z* = -0.05, *p*=.96	β = -6.67 $^{e-04}$, *SE* = 5.55 $^{e-04}$, *Z* =-1.20, *p*=.23	n/a

Homographs (homo)			
Lang x Homo vs Sum	$\beta = 1.61^{e-04}$, $SE = 5.60^{e-04}$, $Z = 0.29$, p =.77	$\beta =-6.41^{e-04}$, $SE = 7.88^{e-04}$, $Z =-0.81$, p = .41	n/a
Lang x homo vs Individual	$\beta =-1.25^{e-04}$, $SE =5.52^{e-04}$, $Z =-0.22$, p=.826	$\beta =-6.92^{e-04}$, $SE = 7.83^{e-04}$, $Z =-0.88$, p=.376,	n/a
Heterographs			
Lang x Het vs Sum	$\beta =-4.35^{e-04}$, $SE = 5.47^{e-04}$, $Z = -0.80$, p=.426	$\beta = -1.00^{e-03}$, $SE = 7.62^{e-04}$, $Z = 1.32$, p =.188	n/a
Lang x het vs Individual	$\beta = 8.12^{e-05}$, $SE = 5.80^{e-04}$, $Z = 0.14$, p =.889	$\beta =-5.35^{e-04}$, SE =, 7.74^{e-04}, $Z = 0.69$, p = 0.49 ,	n/a

*Significance value withstood correction
+Significance did not withstand correction (Holm-bonferroni correction for 48 corrections)
Sum= control group matched to the summed frequency of the control group,
individual= control group matched to the individual frequency of the homophones,
Lexical effect: A tick in this column indicates that there was a significant interaction between the presentation type (immediate or delayed naming) and the effect under consideration, such that the effect was shown in the immediate condition but not in the delayed condition.

Appendix F. A List stimuli for Experiment 2

Homophone		Individual		Summed	
English	Welsh	English	Welsh	English	Welsh
Homographs Matched Controls And Translations					
Iron (Clothes)	Haearn	Leg	Coes	Frog	Llyffant
Lie (Fib)	Gorwedd	Crawl	Cropian	Butter	
Watch (Time)	Oriawr	Spoon	Llwy	Honey	Mel
Nail (Hammer)	Hoelen	Goat	Gafr	Dress	Ffrog
Ring (Finger)	Modrwy	Cloud	Cwmwl	Star	Seren
Bridge (Water)	Pont	Kettle	Tegell	Sleep	Cysgu
Glass (Window)	Gwydr	Wool	Gwlan	Cup	Cwpan
Trunk (Nose)	Truwnc	Shower	Cawod	Milk	Llefrith
Second (First)	Ellliad	Fox	Llwynog	King	Brenin
Bow (Arrow)	Bwa	Cheese	Caws	Laugh	Chwerthin
Letter (Mail)	Llythyr	Blow	Chwythu	Table	Bwrdd
Scale (Measure)	Raddfa	Tree	Coeden	Door	Drws
Ruler (Measure)	Pren Mesur	Neck	Gwddf	Play	Chwarae
Heterographs Matched Controls And Translations					
Flower	Blodyn	Lick	Llyfu	Foot	Troed
Weather	Tywydd	Bell	Cloch	Leaf	Dail
Horse	Ceffyl	Ghost	Ysbryd	Mountain	Mynydd
Meat	Cig	Knee	Glin	Knife	Cyllell
Peace	Heddwch	Kite	Barcud	Tongue	Tafod
Bread	Bara	Lamb	Oen	Drink	Yfed
Weak	Wan	Sword	Cleddyf	Nose	Trwyn
Won	Ennill	Thumb	Bawd	Arm	Braich
Deer	Carw	Bowl	Powlen	Judge	Beirnaid
Shoe	Esgid	Snake	Neidr	Snow	Eira
Bear	Arth	Hammer	Morthwyl	Dog	Ci
Sun	Haul	Stairs	Grisiau	Walk	Cerdded
Two	Dau	Cow	Buwch	Boat	Cwch
Heart	Calon	Finger	Bys	Eat	Bwyta
Knot	Cwlwm	Dream	Breuddwydio	List	Rhestr
Sea	Mor	Chair	Cadair	Bed	Gwely
Bee	Gwenyn	Shoot	Saethu	Sit	Eistedd
Weight	Pwysau	Swim	Nofio	Head	Pen
Hole	Twll	Sheep	Dafad	Run	Rhedeg
Buy	Prynu	Sing	Canu	Hand	Llaw
High	Uchel	Bird	Adar	House	Ty
Tail	Cynffon	Wheel	Olwyn	Man	Gŵr
Sail	Hwyl	Owl	Tylluan	Book	Llyfr
Leek	Cennin	Do	Gwneud	Money	Pres

Appendix G. Mean frequency (log10) and other psycholinguistic variables by experimental groups for the translation task.

		Homophones (n=37)				Homographs (n=13)				Heterographs (n=24)			
		Indiv	Indiv control	Sum	Sum control	Indiv	Indiv control	Sum	Sum control	Indiv	Indiv control	Sum	Sum control
Frequency	Spoken word form	1.22*	0.98*	1.95+	1.81+	1.1	0.78*	1.48	1.67+	1.29*	1.09*	2.20+	1.89+
	Written word form	2.54*	2.53*	3.11+	3.18+	2.44*	2.44*	2.89+	2.99+	2.60*	2.58*	3.24+	3.27+
	Spoken lemma	1.31*	1.08*	2.12+	1.99+	1.27	0.96*	1.68	1.77+	1.34*	1.14*	02.36+	2.10+
	Written Lemma	2.69*	2.63*	3.31+	3.38+	2.65*	2.63*	3.12+	3.11+	2.71*	2.62*	3.41+	3.52+
Syllables		1.16	1.11	1.16	1.14	1.31	1.15	1.31	1.23	1.08	1.08	1.08	1.08
Phonemes		3.24	3.32	3.24	3.41	3.77	3.46	3.77	3.69	2.96	3.25	2.96	3.25
Letters		4.38	4.49	4.32	4.27	4.69	4.54	4.69	4.46	4.21	4.46	4.13	4.17
Ortho N		7.32	5.35	7.27	7.38	6.15	5	6.15	5.46	7.96	5.54	7.88	8.42
Phon N		18.44	15.92	18.44	16.16	13.85	12.62	13.85	13.08	21.04	17.71	21.04	17.83
Translation agreement		79.39	73.31	n/a	79.39	80.05	76.92	n/a	77.88	79.04	73.31	n/a	79.39

*significantly different from the summed homophone group (or summed heterograph/ summed homograph frequency) in a 1-tailed paired t-test (p<.05)
+ significantly different from the individual homophone group(or summed heterograph/ summed homograph frequency) in a 1-tailed paired t-test(p<.05)
Indiv= the frequency of the presented lower frequency homophone partner
Sum= the sum of both homophone partners frequency
Orth N= Number of orthographic neighbours
Phon N= Number of phonological neighbours

Appendix H. The mean reaction times (standard deviations) for the translation study

	Homograph	Heterograph
Homophone	1206.27(448.01)	1059.54 (384.39)
Individual control	1141.95(423.04)	1184.90 (448.68)
Summed control	1115.49(435.68)	1004.07 (355.52)

Appendix I. The various models tested with the *p* values and chi squared (Chisq) values from ANOVAs comparing each model with the previously significant model using the translation inverted RTS

	Model	Chisq	p
1	(1|subject)+(1|items)		
2	(condition)+(1|subject)+(1|item)	13.652	<.001
3	(Spelling+condition)+(1|subject)+(1|item)	2.82	0.09
4	(Spelling*condition)+(1|subject)+(1|item)	6.67	0.08
5	(Spelling*condition)+ Name agreement+(1|subject)+(1|item)	24.55	<.001
6	(Spelling*condition)+ name agreement+ phon N +(1| subject)+(1|item)	2.48	0.12
7	(Spelling*condition)+ name agreement+ number of phonemes+(1|subject)+(1|item)	0.14	0.71
8	(Spelling*condition)+ name agreement preceding rt+(1| subject)+(1|item)	10.75	<.001
9	(Spelling*condition)+ name agreement +preceding rt +presentation order +(1| subject)+(1|item)	13.41	<.001

+ despite the spelling condition and the spelling*condition only trended towards improving the model, they were included as there was strong apirori reasons to. The main effect of spelling is nested within the fixed and would not influence RTs as a simple main effect. Secondly the interaction needs to be present in the lme in order for perform contrasts and disentangle these nested main effects. Therefore it is included in subsequent competitive modelling.
Spelling= heteterograph, homographLanguage- monolingual, bilingual
Condition= homophone, individual frequency matched control, summed frequency matched control
Nameagreement= percentage of correct name produced
Phoneme= number of phonemes
Phon N= number of phonological neighbours
preceedingRT= the latency of the preceding item
Original= size of the image (a visual complexity measure)
Order= the presentation order

Chapter 3

Too harts won sole: Dysgraphia Treatment Outcome as Evidence for Homophone Representation

There is general agreement that each word in our language requires a stored representation of meaning (semantics), and phonological/orthographic form(e.g., Caramazza, 1997; Dell, 1986; Levelt, Roelofs, & Meyer, 1999). However, there are circumstances when a word's form completely overlaps with that of another word (e.g., *cricket* the insect, and *cricket* the sport). These types of words (homophones) clearly have separate meanings and hence separate semantic representations but perhaps their form need not be separately represented? Homophone representation has been investigated over the past 20 years using various different techniques (e.g., frequency effects on reaction times, neuropsychological treatment, error rates) yet it is still unclear how homophones are represented in the lexicon. In this study, we explore the orthographic representation of homophones and their written production.

Representations in the lexicon

There are currently no models that focus solely on orthographic homophone representation. In contrast, there are three competing models of homophone representation in spoken production, which are useful for conceptualising written homophone production given the lack of orthographic production models. Figure 1 depicts these three theories of spoken homophone production, whereby in (1a) homophones share a lexical representation as in, for example, Levelt et al.'s Two Stage model (1999; Figure 1a). According to (1b) homophones have independent representations with strictly feed forward activation (Figure 1b; Caramazza, Costa, Miozzo, & Bi, 2001), while in (1c) homophones have independent representations with interactive activation between the lexical representations and segment level (Figure 1c; Middleton, Chen & Verkuilen's , 2015 adaptation of Schwartz, Dell, Martin, Gahl, & Sobel, 2006). Due to similarities between spoken and written production (Damien) it seems plausible that phonological and orthographic architecture are similar. Therefore we can assume that the orthographic representation of homophones is one of the figures in Figure 1, however the lexical representation would be orthographic (not phonological) and the segment level would be graphemes (not phonemes). Previous psycholinguistic studies have offered competing evidence for the different types of models shown in Figure 1 (e.g., Caramazza et al., 2001; Dell, 1990; Jescheniak & Levelt, 1994), however here we focus on neuropsychological evidence.

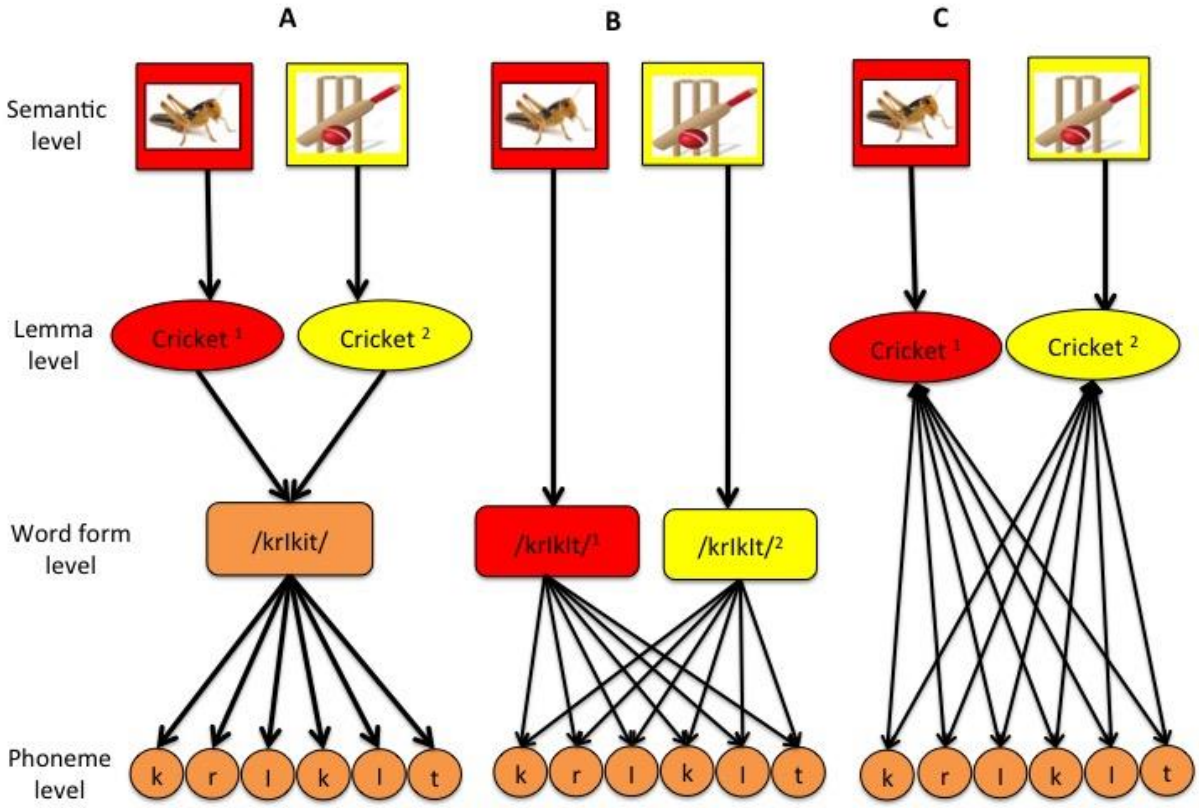


Figure 1: Three architectures for homophone representation in the lexicon a) separate lemmas and shared (Levelt et al., 1999) b) no lemma level and separate modality specific word forms (Caramazza et al., 2001) and c) lemmas but no word form representations and interactive activation (Middleton et al., 2015; Schwartz et al., 2006).

Treatment studies of spoken production

One method of investigating homophone representation involves exploring generalisation of the effects of language therapy with people with aphasia (PWA). It is assumed that Previous studies have found that treating one homophone (e.g., *knight*) improves spoken naming of the untreated homophone partner (e.g., *night*; Biedermann, Blanken, & Nickels, 2002; Biedermann & Nickels, 2008a, 2008b; Biran, Gvion, Sharabi, & Gil, 2013; Blanken, 1989). Following Blanken's (1989) landmark study, Biedermann and colleagues replicated the pattern of homophone generalisation in both German (one person) and English PWA (two people) with lexical-access impairments. Biedermann and Nickels (2008a) found this effect for both homographic homophones (e.g., *cricket* [insect]/*cricket* [sport]) and heterographic homophones (e.g., *knight/night*), with no difference in the extent of generalisation explained by orthographic similarity. However, there was no generalisation to phonologically related items, therefore, the authors interpreted generalisation as being due to *shared* representations at the phonological word form level (support for Figure 1a).

Biran et al. (2013) replicated this work in Hebrew using a phonological cueing therapy with two PWA who also had word retrieval deficits. Both participants showed improvement in naming treated homographic homophones (e.g., *mapa*; tablecloth) and their untreated homophone partners (e.g., *mapa*; map), but not phonologically related controls (e.g., *maca;* Matzah).

In sum, to date, every phonological treatment study that has addressed spoken homophone production has found generalisation of treatment to untreated homophones. Furthermore, the effects have been interpreted as evidence for shared representations at the phonological word form (Biedermann et al., 2002; Biedermann & Nickels, 2008a, 2008b; Biran et al., 2013). These findings support the theory that Levelt et al. (1999) propose in their Two-Stage Model of spoken language production (Figure 1a); shared phonological word forms for homophones. This is also supported by data from homophone translation (Jescheniak & Levelt, 1994; Jescheniak et al., 2003).

However an alternative explanation for the homophone generalisation effects described above, is that generalisation is caused by feedback from phonemes to independent representations (as in Figure 1c). Treatment of one word form would result in increased activation of the target's phonemes. Activation from these phonemes could feedback and activate other word forms that also contain these phonemes, therefore the word form with the greatest amount of phonolgoical overlap would show the most generalisation, with homophones showing greatest generalisation given their 100% overlap. Studies examining orthographic treatment in PWA have attributed generalisation to orthographically related items (neighbours) to the orthographic equivalent of this feedback (e.g., Harris et al., 2012; Sage & Ellis, 2006; see below). In their study, Biedermann and Nickels (2008b) rejected this feedback account as they found no evidence for differing generalisation to phonologically related non-homophones dependent on number of shared phonemes (there was no generalisation at all). However, as the study was not explicitly designed to investigate generalisation depending on varying phonological overlap, the amount of variation in phonological overlap was not adequate to thoroughly investigate this. Therefore, it still remains debatable whether homophone generalisation is caused by a shared word form or by feedback to independent word form representations. This, amongst other reasons, leads us to study generalisation in the orthographic lexicon in order to disentangle homophone representation.

Treatment studies of written production

It is well established that there are separate and independent (but related) orthographic and phonological output lexicons (e.g., Shelton and Weinrich, 1997). However, as noted above, most research investigating the production of homophones has involved the spoken modality- indeed few models address the representation of written homophones at all (but see Caramazza, 1997). Therefore, in order to gain a comprehensive understanding of homophone representation in the language system, it is critical to investigate whether the orthographic lexicon is structured similarly to the phonological lexicon. If homophones share a phonological word form, do they also share an orthographic word form? To our knowledge, there has been only one previous study investigating orthographic treatment of written homophone naming. Behrmann (1987) conducted a remediation study with a participant (CCM) who had impaired access to the orthographic output lexicon resulting in poor irregular spelling with phonologically plausible errors and homophone confusions. Before treatment, CCM scored 49% on homophone spelling to dictation (with a disambiguating sentence) and a large proportion (57%) of errors were the homophone partner (e.g. writing 'sail for 'sale'). On this basis, it was hypothesised that CCM had difficulty retrieving the homophone spelling on the basis of semantic information (accessing the orthographic output lexicon from semantics). CCM was treated using a series of tasks that involved contrasting the different spellings of homophones, using pictures and written sentence completion tasks (e.g., she was shown a picture of a knight, with the written word *knight*, which was orthographically contrasted with the word *night*, and then completed a written sentence requiring the target spelling *knight*). The treatment improved homophone spelling of the treated homophones and this generalised to better spelling of untreated irregular words but not homophone partners. Behrmann (1987) suggested this was due to improved lexical access (for the treated homophones) as well as an improved visual-checking mechanism (generalisation to irregular words). This was supported by the pattern of responses to untreated homophone partner: the percentage of non-word errors was greatly reduced (from 36% to 18%) and a larger percentage of errors were the (treated) homophone partner (82%) compared to pre treatment (57%).

Behrmann's (1987) treatment investigated written production of heterographic homophones. Although previous spoken homophone production treatment found no difference in generalisation for heterographic or homographic homophones, critically both types of homophone share phonology (but differ in orthography) and the treatment focused on improving access to phonological forms. In contrast, Behrmann's (1987) treatment focused on improving access to orthography for homophones that are orthographically different , hence perhaps it is unsurprising that there was no homophone generalisation (instead the irregular word generalisation was due to non-lexical strategies). It remains to be investigated whether treatment of access to orthography for words that share orthography (and phonology), that is homographic homophones, results in generalisation. There are also some methodological weakness in Behrmann's (1987) treatment design which could have led to reduced homophone generalisation. The most notable being that the selection of items for treatment may have biased against finding generalisation as untreated stimuli were closer to ceiling (i.e., the homophones selected for treatment were those that were incorrect (0/50) whereas the untreated homophones were spelled correctly (30/50).

There are other instances of differing patterns of generalisation across phonological and orthographic treatment modalities. For example, while there has been a lack of generalisation to phonological neighbours (e.g. cricket-ticket) in some spoken treatment (e.g., Biedermann & Nickels, 2008b), generalisation to orthographic neighbours (e.g. clock-block) has been found in some written treatment (e.g., Harris et al., 2012).

Harris et al. (2012) and Sage and Ellis (2006) both conducted treatment studies with individuals with acquired impairments at the graphemic buffer (the level of temporary storage of graphemes before processing for output) and found that treatment generalised to untreated direct neighbours (pairs of words with one grapheme different). The authors suggested this generalisation could have been caused by the treatment increasing activation of the orthographic representations of the treated items (e.g. clock), which in turn activated the graphemes stored in the graphemic buffer (e.g. c-l-o-c-k). These activated graphemes would have fed activation back to the word forms that shared these letters (e.g., the graphemes l-o-c-k- would feed activation back to 'block') resulting in subsequent improved production of both the target and neighbours. However Krajenbrink, Nickels and Kohnen (2016) failed to replicate this direct neighbour generalisation in two similar cases of graphemic buffer impairment. One possible explanation was that, in these individuals, perhaps there was reduced or absent feedback from the graphemic buffer to the lexicon (Krajenbrink et al., 2016).

Generalisation from treated to untreated spelling has also been found in cases of developmental dysgraphia with no graphemic buffer impairment (e.g., Brunsdon, Coltheart, & Nickels, 2005; Kohnen, Nickels, Coltheart, & Brunsdon, 2008). Kohnen et al. (2008) found that this generalisation was dependent on orthographic neighbourhood. When a word is spelled similarly to many of other words it has a high orthographic neighbourhood (e.g., line has 22 'neighbours', lime, lane, lint, pine etc), whereas a word has a low orthographic neighbourhood if no word is one letter different in its spelling (e.g., church has no neighbours). Similar to the studies with acquired dysgraphia, Kohnen et al. (2008) suggested that untreated words with high orthographic neighbourhoods are more likely to improve due to increased feedback from graphemes to orthographic word forms that are repeatedly activated due to the high amount of shared graphemes with frequently produced words.

This feedback mechanism, as the cause of generalisation to direct neighbours in written production, is analogous to the alternative (feedback) explanation for generalisation to untreated homophones in spoken production mentioned earlier (Biedermann & Nickels, 2008b, 2008a). However if it were the case that generalisation in spoken modality was due to feedback, there should have also been generalisation to phonological neighbours (which was not found), as was found to orthographic neighbours (e.g., Harris et al., 2012).

In sum, it is still unclear whether the effects of generalisation found for phonological homophone treatment are due to feedback or shared-word forms, nor is the nature of the orthographic representation of homophones clear. Therefore, we carried out a treatment study that aimed to further investigate whether any generalisation in orthographic homophone treatment is due to improved shared lexical entries, or feedback to separate entries and how this differs across heterographic and homographic homophone spelling. We thus treated three groups of stimuli (homographs, heterographs, non-homophonic controls) and investigated generalisation to five untreated groups (homograph partners, heterograph partners, direct neighbours of the controls, unrelated high orthographic neighbourhood words and unrelated low orthographic words).

If the mechanism for generalisation is feedback from graphemes, then we would expect to find generalisation to homographs plus (less) generalisation to items that have orthographic overlap (heterographs and direct neighbours), as in Figure 1c. If we find generalisation only for homographic homophones this will suggest they share a word form representation and there is no feedback from graphemes, as in Figure 1a. If no generalisation is found, this is consistent with homographs having separate representations with strictly feed forward activation, as in Figure 1b.

Case History

The participant in this study, CWS, was a 67-year-old right-handed, former builder from North Wales. He learnt both Welsh and English before the age of six and used both regularly. Both pre- and post-stroke he was equally proficient in English and Welsh, however, as this treatment investigates English, only his English naming performance is reported (see Roberts, 2013, for a comprehensive report of his bilingual language abilities). CWS suffered a right frontal infarct in 1997 (18 years prior to this experiment). This resulted in left-sided hemiplegia and crossed-aphasia (aphasia due to right hemisphere damage despite right-handedness) resulting in agrammatic, non-fluent speech. Table 1 shows CWS's language performance on a range of standardised (to English monolinguals) tests. CWS's spoken and written comprehension remains intact along with word and non-word repetition. CWS performed within the control range for spoken object and action naming (Druks & Masterson, 2000). His visual word recognition was just below ceiling and within control performance range. Regular and irregular reading aloud were intact, although he showed severely impaired non-word reading, a symptom pattern that is consistent with phonological dyslexia (see Tainturier, Roberts, & Leek, 2011, for detailed analysis of his reading).

Table1. English Language background assessment for CWS

Task	Number of items	CWS correct %	Control Mean %	Control Min. %	Control SD %	N of Controls
Comprehension						
PALPA 47 oral word-picture matching	40	100%	98	87	2.67	31
PALPA 48 written word-picture matching	40	100%	99	87	1.53	32
Pyramids and Palm Trees Test; three pictures	52	96%	98	94	-	13
Single word Repetition						
Word repetition (Bangor University)	80	100%	-	-	-	-
English Non-word repetition (Bangor University)	40	100%	-	-	-	-
Spoken Naming						
Object Naming battery (list B)	81	91%	-	91	-	40
Action Naming battery (list B)	50	86%	-	86	-	40
Visual Word Recognition						
PALPA 25: Real and non-words lexical decision[++]	120	98%	99	-	0.54	26
PALPA 3 minimal pairs: Written word selection	72	97%	97	-	2.35	23
Reading						
PALPA 19 upper case to lower case letter matching	26	96%	100	96	0.77	26
PALPA 32 grammatical class reading	**80**	95%*	100	-	1.45	32
PALPA 34 lexical morphology and reading	**90**	90%*	-	-		-
English Reading: Regular words (Bangor University) [+]	40	98%	99	95	1.65	20
English Reading: Irregular words (Bangor University)[+]	40	100%	99	95	1.73	20
English Reading: Non-words (Bangor University)[+]	**40**	**43%****	95	83	5.10	20
PALPA 53: Cross Modality Comparisons						
Oral picture naming	**40**	**65%****	99	-	0.87	29
Irregular	20	80%	-	-	-	-
Regular	20	60%	-	-	-	-
Written picture naming	**40**	**15%****	97		3.33	27
Irregular	20	5%	-	-	-	-
Regular	20	25%	-	-	-	-
Spelling to dictation	**40**	**7.5%****	99	-	-	2
Irregular	20	0%*	-	-	-	-
Regular	20	15%*	-	-	-	-
Repetition	40	97.5%	99	-	2.05	28
Irregular	20	100%	-	-	-	-
Regular	20	95%	-	-	-	-

+Control scores taken from aged matched control monolingual participants from Bangor University. All other control data is from the appropriate published test.
++ Average mean and standard deviations across the subsets
Bold represents scores which are impaired (2.5 standard deviations below control mean)
** Scores that are at least two standard deviations below control mean.
* Scores that are thought to be impaired to some degree, but normative data is not available.

Table 2. In-depth spelling-to-dictation assessments

Task	Number of stimuli	CWS % correct	Control Mean %	Control Range %	Control SD %	Control N
Regularity *(Bangor University)*						
Regular words	80+	**41**	97	80-100	5.23	20
Irregular words	80+	**18**	91	55-100	12.40	20
Non-words	80+	**39**	74	53-93	12.07	20
Frequency (Words collapsed across JHU# lists)						
High-frequency words	147	21*	-	-	-	-
Low-Frequency words	146	12*	-	-	-	-
Length (JHU list length)						
4-5 Letters	27	30*	99	-	-	5
6 letters	15	7*	92	-	-	5
7+ letters	28	7*	93	-	-	5
Grammatical Category (JHU part-of-speech)						
Nouns	28	4*	-	-	-	-
Verbs	28	4*	-	-	-	-
Adjectives	28	7*	-	-	-	-
Nonwords	34	12*	-	-	-	-
Concreteness (JHU)						
Concrete words	21	19*	98	-	-	5
Abstract words	21	0*	91	-	-	5
Copy						
Direct copy (PALPA 44)	40	98	-	-	-	-
Delayed copy transcoding						
Regular words	20	80*	-	-	-	-
Irregular words	20	45*	-	-	-	-

#JHU= John Hopkins University Dysgraphia Battery (Goodman & Caramazza, 1985)
+ Double administration for CWS, therefore control number of items = 40
Impaired Scores in bold (2.5 standard deviations below control mean)
* Scores that are thought to be impaired to some degree, but normative data is not available.

CWS's spelling performance is reported in further detail in Table 2. Although CWS was impaired in all aspects of spelling, he had a significantly larger deficit when spelling irregular words compared to regular words and non-words (Chi squared, $X^2(1, N=80)= 12.99$, $p<.001$, $X^2(1, N=80)= 10.90$, $p=.001$). During written picture naming tasks (e.g. PALPA 53), despite being asked not to name or spell aloud, CWS would attempt to spell by breaking a word down into phonemes (which he could do accurately) and spelling one phoneme at a time. This pattern of performance suggests attempted use of a sub-lexical strategy secondary to damage to the orthographic output lexicon or access to this lexicon. This strategy results in better performance on regular word spelling compared to irregular word spelling and phonologically plausible (regularisation) errors. Table 3 shows examples of errors taken from Roberts (2013), indeed 39% of CWS's spelling errors were phonologically plausible. However, CWS also produced a large number of phonologically implausible errors (33%). To summarise, CWS presented with a mixed dysgraphia profile. He had a clear lexical impairment, however the presence of a length effect and impaired nonword spelling also suggests possible additional graphemic buffer and/or sub-lexical impairments (for a detailed analysis of this see Roberts; 2013).

Table 3. Errors made by CWS in spelling to dictation taken from Roberts (2013)

Error type	**Example**	**Words % (N= 599)**	**Non-words % (N=45)**
Phonologically plausible errors	Into-> INTU	**39**	1
Real word error	Work->WORD	**7**	36
Phonologically implausible nonwords (50% or more letters correct)	Hotel->HOTOL	**33**	7
Phonologically implausible nonwords (less than 50% target letters correct)	Feather->FAFARA	**8**	1
Cross language errors	Nine->NAIN	**12.52**	2.22

CWS's orthographic word form level impairment made him a suitable candidate to investigate generalisation of homophone treatment at the orthographic word form level.

Intervention Study

The study comprised two treatment phases, which are reported in turn, below.

Phase 1

Phase 1 consisted of a copy and recall treatment (CART) (e.g. Beeson, 1999) in the presence of the picture. This task was chosen because we wanted to ensure that the treatment did not improve spelling via phoneme to grapheme conversions. Instead increased the accessibility of the orthographic lexicon (as this was the representation we were investigating). Previous treatment studies have found CART to be an effective strategy to strengthen orthographic representations (Beeson, Hirsch, & Rewega, 2002).

Method: Phase 1

Stimuli.

Stimuli were picturable nouns presented as photographs 300 x 300 pixels in size, displayed in the centre of a computer screen with written descriptions underneath. Descriptions were designed to clarify picture identity to facilitate written naming. They were not 'stand-alone descriptions' and did not contain any target semantic competitors (e.g. the definition used with knife was 'used for eating'). All stimuli (picture with the description) had over 70% name agreement from 10 control participants (mean age 29.20 years).

Stimuli were assigned to one of eight experimental subsets: 1) Homographic homophones (e.g., *cricket* [insect]), (2) homographic partners of 1 (e.g., *cricket* [sports]), (3) Heterographic homophones (e.g., *sale*), (4) heterographic partners of 3 (e.g., *sail*), (5) Non-homophonic controls with direct neighbours (e.g., *bath*), (6) direct orthographic neighbours of 5 (e.g., *path*), (7) non-homophonic control words with high orthographic neighbourhoods (e.g., *line*) (8), and non-homophonic control words with very low orthographic neighbourhoods (e.g., *church*). The direct neighbours in subsets 5 and 6 consisted of words with one grapheme substituted (instead they differed by one grapheme in the same position e.g., *cake-cave*; (Coltheart, Davelaar, Jonasson, & Besner, 1977). They did not include additions or substitutions. Homophone and words with direct neighbours (subsets 1-6) were randomly assigned (using the Excel RAND function) to two sets (per condition i.e. 1 or 2, 3 or 4, 5 or 6). These sets were then adjusted to ensure matching on the variables presented in Table 4 before randomly being assigned to treated or untreated conditions. All variables bar frequency and regularity were obtained from N-watch (Davis, 2005). CELEX (Baayen, Piepenbrock, & Gulikers, 1996) was used to obtain Frequency counts per million, then log10 transformations were performed. An item was listed as irregular if it had at least one grapheme that was classed as not regular by Perry, Ziegler and Coltheart (2002) or rare by Fry (2004).

Table 4. Matching of experimental subsets on accuracy, log frequency and other psycholinguistic variables for Phase 1.

	Homographs		Heterographs		Non-homophonic controls			
Subset	1	2	3	4	5	6	7	8
	Treated (N=22)	Untreated partners (N=22)	Treated (N= 14)	Untreated partners (N=14)	Treated (N=24)	Untreated Direct neighbours of 5 (N=24)	Untreated High ON (N=26)	Untreated Low ON (N=22)
Written accuracy Baseline 1	9.09	9.09	14.29	7.14	16.67	16.67	25.93	4.55
Written accuracy Baseline 2	13.64	9.09	21.43	21.43	25.00	16.67	22.22	4.55
Written accuracy Baseline 3	0.00	4.55	14.29	14.29	12.50	8.33	23.63	4.55
Spoken accuracy Baseline 1	95.45	77.27	100.00	85.71	83.33	79.17	88.89	72.73
Spoken accuracy Baseline 2	59.09	63.64	64.29	85.71	70.83	70.83	77.78	72.73
Spoken accuracy Baseline 3	81.82	72.73	92.86	71.43	66.67	75.00	81.48	68.81
Frequency: written lemma (log10)	2.06	2.20	2.41	2.62	2.67	2.76	2.86	2.67
Frequency: written word form (log10)	0.57	0.71	1.06	1.19	1.05	0.96	1.15	0.85
Frequency: spoken lemma (log10)	0.69	0.84	1.17	1.29	1.17	1.14	1.25	1.00
Frequency: spoken word form (log10)	1.88	2.01	2.23	2.50	2.48	2.59	2.74	2.49
Syllables	1.14	1.14	1.07	1.07	1.00	1.00	1.00	1.00
Phonemes	3.64	3.64	3.00	3.00	3.09	3.25	3.38	3.19
Letters	4.32	4.32	4.21	4.36	3.95	4.00	3.88	5.05*
Orthographic neighbourhood density	7.41	7.41	7.36	8.57	10.91	10.21	8.62	0.57*
Orthographic neighbourhood freq.	106.61	106.61	62.27	385.25	86.53	129.83	310.67	27.82*
Phonological neighbourhood density	16.05	16.05	21.29	21.29	21.09	20.50	14.92*	11.62*
Phonological neighbourhood freq.	86.14	86.14	451.21	451.21	184.22	306.20	279.22	286.50
Regularity	0.91	0.91	0.57	0.50	0.92	0.92	0.81	0.55*

*Significantly different from matched subset
Orthographic neighbourhood density = number of words with one letter difference
Orthographic neighbourhood freq. = average frequency of all the orthographic neighbours
Phonological neighbourhood density = number of words with one phoneme difference
Phonological neighbourhood freq.= average frequency of all the phonological neighbours

The treated and untreated subsets were matched for accuracy across the three baseline sessions (see Table 4). The treated and untreated paired subsets were matched for all the psycholinguistic variables shown in Table 4 (e.g., subset 1 had the same frequency as subset 2), which were obtained from N-watch (Davis, 2005). The untreated orthographic neighbourhood sets (subsets 7 and 8) were also matched to subset 5 (and therefore also 6) on all variables except the orthographic neighbourhood variables (or any that are correlated to this) for subset 8. As the low orthographic neighbourhood control subset (subset 8), was designed to have lower orthographic neighbourhood than subset 7 (and therefore 5 and 6) any variables associated with orthographic neighbourhood were significantly different from these subsets. None of the subsets had any direct substitution neighbours in another subset apart from subset 5 that had exclusively direct neighbours in subset 6.

Procedure.

All subsets were presented for picture naming over three pre-treatment baseline tests and three post-tests. Only stimuli from subsets 1, 3 and 5 were treated. Generalisation to 2, 4, 6, 7 and 8 was investigated by comparing baselines to post-tests.

A timeline is presented in Figure 2. Both written and spoken word production were tested during baselines, once within the treatment phase and at post-tests. All assessments were two weeks apart, except for the final two post-tests. The experimenter read out the description as the picture was presented. Each session presented the items in a different random order.

Spoken naming familiarisation.

Two days prior to baseline testing, two sessions of spoken naming familiarisation were conducted, two days apart from each other. This was to ensure that CWS was familiar with the picture names, to rule out any incorrect answers due to ambiguous pictures and to ensure that the phonology of each item was equally available. The spoken familiarisation phase consisted of presentation of the stimulus picture and spoken name for CWS to repeat. The stimuli were split evenly into two sets. Each set contained During Session 1, Set 1 was presented first with the correct name for CWS to repeat. The same items were subsequently presented for CWS to name, if he was correct he was given feedback (e.g. 'well done that is correct'). If he produced the wrong name or no response the correct name was given for repetition (which was always correct). Set 2 was then presented using the same procedure. Session 2 consisted of the same procedure as Session 1 however the session began with Set 2 such that, in total each item was repeated and named four times (twice in each session).

Assessment sessions.

Because of the large number of items, each assessment was split into two sessions. Each set contained 84 experimental items and 16 filler items (which were included for data collection for a separate study). During the first session of Baseline 1, CWS wrote the names of Set 1, and completed spoken naming of the pictures of Set 2. During the second session of Baseline 1 CWS was asked to write the names of Set 2 and complete spoken naming of Set 1. Whether Set 1 or Set 2 was given for written naming in the first or second session of an assessment was alternated across time points. The two sessions assessing both modalities of one set were at least one day apart. As we were primarily interested in investigating the effect of treatment on written naming performance, and only secondarily in the effect of spelling treatment on spoken naming (which was close to ceiling), all of our baselines and post-tests assessed written naming before spoken naming. Each set was given in a different randomised order at each presentation time point.

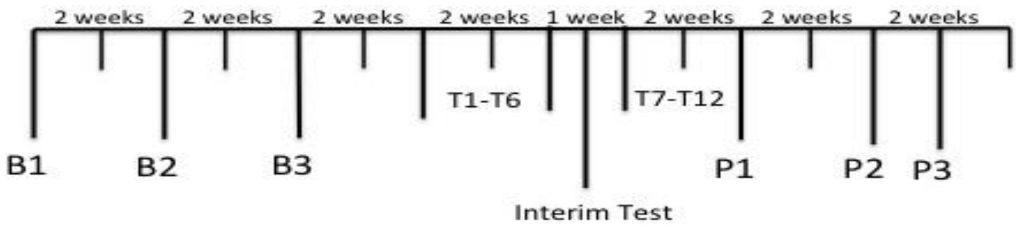


Figure 2: Time line of baselines (B1, B2, B3), treatment sessions (T1-12) and post-tests (P1, P2, P3).

Treatment consisted of six sessions over two weeks before an interim assessment and then another six sessions over two weeks. Overall, CWS received 12 treatment sessions over four weeks. Each treatment session contained 60 experimental items (i.e. one of the homograph, heterograph and direct neighbour partners) as well as 15 filler items. CWS also completed six homework sessions, one after every second treatment session. The first post-test occurred two days after the last treatment session, then, Post-test 2 was two weeks later, followed a week later by Post-test 3.

Treatment.

Treatment was based on copy and recall (CART) approach from Beeson (1999). CWS was presented with a stimulus picture and correct spelling of the target name and asked to copy the word while the word stayed in sight. This immediate copying was excellent, and this stage of the treatment, was error-free. The experimenter then covered both the presented correct spelling and CWS's immediate copy and counted to ten aloud. After this ten second delay CWS was asked to 'try and spell the name of the picture again'. The correct spelling was then presented, and CWS was asked to judge whether he had correctly spelled the target item. If he had (correctly) confirmed his delayed copy was correctly spelled, feedback was given and the next item was presented. If he misspelled the item and realised this was the case, he was asked what part he thought was wrong. Then the target word was presented and contrasted to his incorrect spelling, before he copied it once again. On the few occasions he incorrectly judged his spelling as correct the experimenter presented the target item and pointed out the contrast, before he copied the correct target. On the rare occasion he was unable to produce anything from delayed copying the target was presented again to copy. Therefore, each item was written correctly twice in each treatment session. Homework also consisted of immediate copying and delayed recall. CWS was given a booklet with one picture per page – on one page the written word was presented below the stimulus picture for copying. This would be followed by another stimulus item for copying, and then the first picture presented for delayed recall after this intervening item, followed by the second item for delayed recall.

Analysis.

We compared pre- and post-treatment scores using Weighted Statistics (Howard, Best, & Nickels, 2015). We initially established if there was overall improvement over the course of the study by conducting a trend (WEST-Trend) analysis. If there was significant improvement, then we also conducted a Rate of Change (WEST-ROC) analysis to investigate if this trend could be attributed to improvement during the treatment phase. Only if both these tests were significant did we conclude there was treatment-specific improvement (Howard et al., 2015). For written naming we analysed both whole word accuracy and letter accuracy using letter accuracy scoring adapted from Buchwald and Rapp (2009). Each letter attracted a score of one if it was correctly produced in the correct position. Between 0.25-1 points were deducted for each letter position that was either transposed, migrated, substituted, missed or included an additional letter. We did not count the same letter position twice; if two errors fell on the same letter position, we only scored one[6].

Results: Phase 1

We only report the spelling analysis below. As our primary interest is on written naming, the results and analysis of spoken naming are presented in Appendix A (Table 1). In brief however, spoken naming was close to ceiling (as was intended by the familiarisation phase) and there was no statistically significant treatment-related change.

Due to unforeseeable circumstances in Phase 2[7], it was only possible to obtain one post-test. Therefore, in order to compare the two phases, Phase 1 analysis with only one post-test is reported below (see Table 5 and Figures 3-5), with Phase 1 analysis with three post-tests is in Appendix A (Table A2 and A3). All the effects were in the same direction across analysis with one and three post-tests; however, improvement was not maintained for all items.

[6] For example, when scoring KNIFE spelled as *neafh* the E has moved from fifth position to the second, so it was scored 0.5. There is the addition of H after the F however this does not mean the fifth position loses another 0.5 to be 0, instead the score is capped at 0.5

[7] CWS was admitted to hospital with pneumonia immediately after Post-test 1.

The treated items overall (treated homographs, heterographs and non-homophonic controls) improved due to treatment for both the whole word and letter accuracy (see Figure 3). The related but untreated (homographic partners, heterographic partners, and direct neighbours) also significantly improved whereas the untreated unrelated items (untreated controls high and low in orthographic neighbourhood) did not improve over the course of the study. The treated items improved significantly more than the untreated unrelated items (two sample t-test, t= 4.50, *p*<.001) and the untreated related items (two-sample t-test, t=4.52, *p*<.001).

Given the clear treatment-related improvement on the treated items, we therefore, looked further into the effects of treatment on the different experimental subsets.

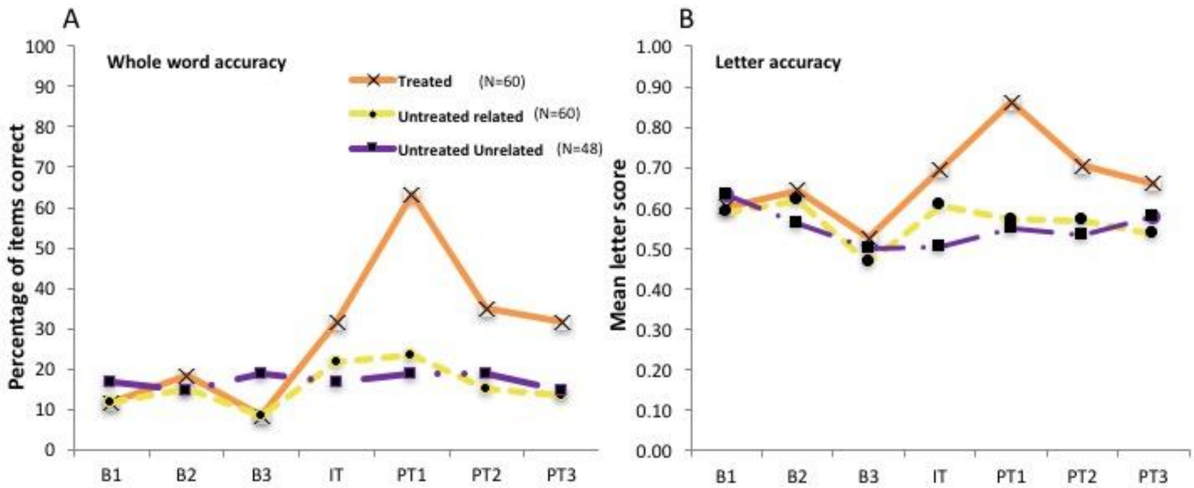


Figure 3: Phase 1 performance scored by a) whole word accuracy, and b) letter accuracy overall letter scoring for the combined treated items (homographs, heterographs and non-homophonic controls), untreated related (homograph and heterograph partners and direct neighbours of non-homophonic controls) and untreated unrelated (high and Low orthographic neighbourhood controls)

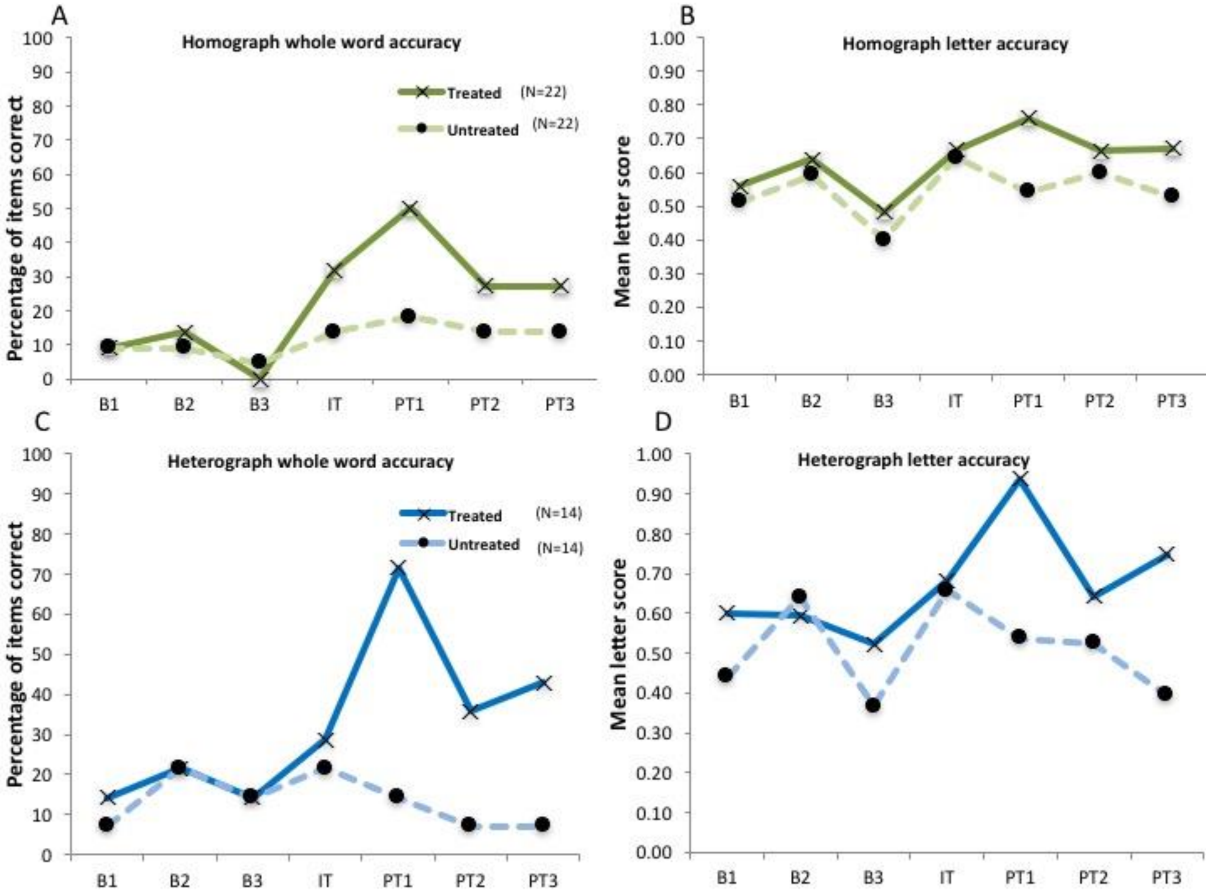


Figure 4: Phase 1 accuracy for homographs (top panels) and heterographs (bottom panels) on whole word accuracy (A&C), and letter accuracy (B&D)

There was significant improvement due to treatment both for word and letter accuracy for the treated homographs (Figure 4a & b) and treated heterographs (Figure 4c & d). However, there was no treatment-related improvement for the untreated homographs or heterographs either word or letter accuracy.

The treated non-homophonic controls also improved due to treatment in both whole word and letter analyses (as shown in Figure 5a & b and Table 5). While, in the whole word analysis, the untreated direct neighbours of controls did show significant improvement, this was not replicated across the letter level analysis. There was no treatment-related improvement for whole word or letter accuracy for either orthographic neighbourhood control sets (Figure5c & d, Table 5).

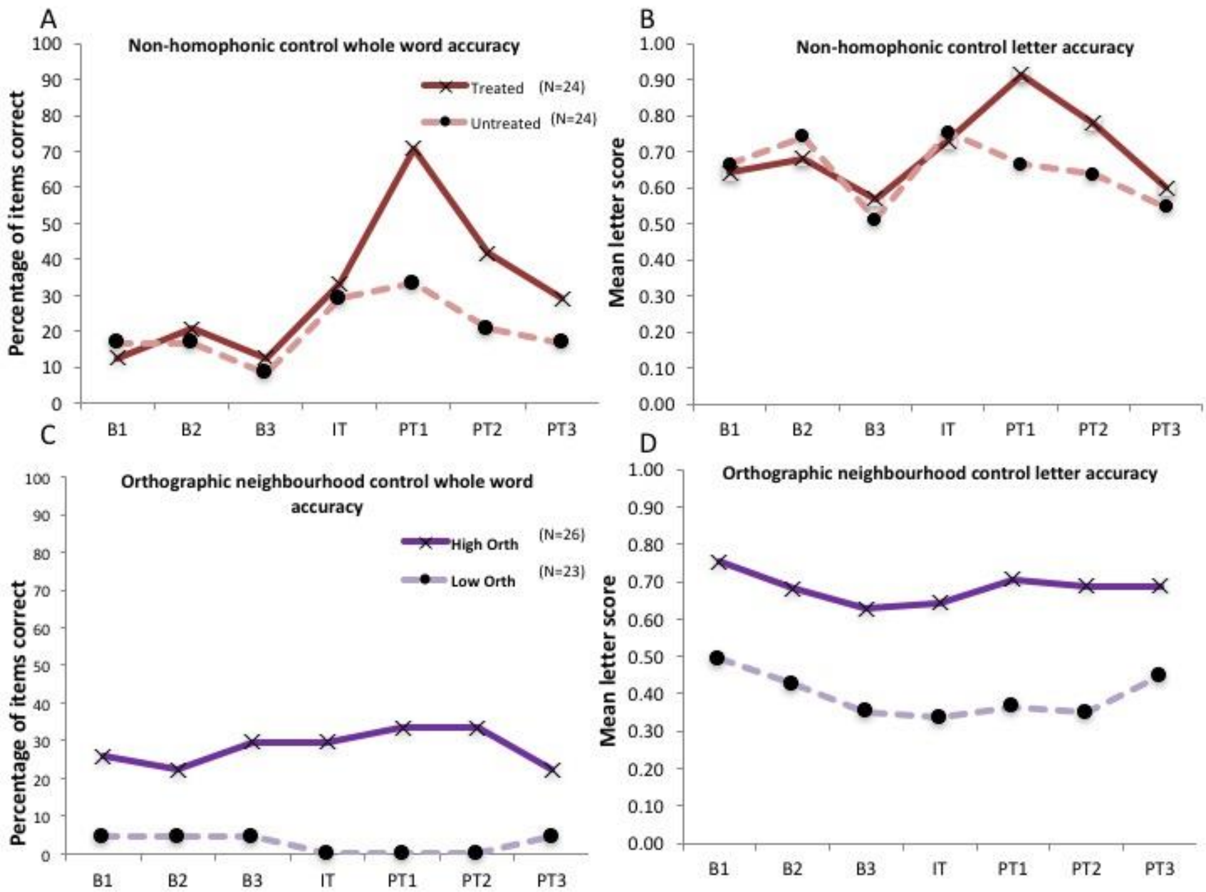


Figure 5: Phase 1 accuracy for non-homophonic controls and direct neighbours (top row) and orthographic neighbourhood controls (bottom row) on whole word accuracy (A&B), and letter accuracy (B&D).

Table 5. Results of the WEST t-tests and their significance for Phase 1 of the treatment across the three baselines, interim test and first post-test.

	Phase 1 (one post test)						
	Whole word accuracy			Letter accuracy			
Subset (degrees of freedom)	West trend	West-ROC	Treatment specific improvement	West trend	West-ROC	Treatment specific improvement	Consistent
Treated (1,59)	6.25**	6.47 **	✓	5.22 **	6.50 **	✓	yes
Untreated related (1,59)	2.37*	2.24 *	✓	0.10		✗	no
Untreated unrelated (1,47)	0.43		✗	-2.11		✗	yes
Subsets							
Homographs:							
1.Treated (1,21)	3.24 **	4.24 **	✓	2.02 *	2.94 **	✓	yes
2.Untreated (1,21)	1.23		✗	0.53		✗	yes
Heterographs:							
3.Treated (1,13)	3.32 **	2.73 *	✓	3.64 **	4.00 **	✓	yes
4.Untreated (1,13)	0.46			1.66		✗	yes
Non-homophone:							
5.Treated (1,23)	4.92 **	3.94 **	✓	3.89 **	3.65**	✓	yes
6.Direct N (1,23)	2.31 *	2.04 *	✓	0.09			no
Unrelated Untreated:							
7.High ON (1,25)	1.32		✗	-0.91		✗	yes
8.Low ON (1,22)	-1.00		✗	-2.14		✗	yes

*$p<.05$ **$p<.001$
Consistent= Consistent across both types of analysis (word and letter accuracy)
✓ = both West-trend and West-roc were significant indicating significant improvement over the course of the study and greater improvement during the treatment than no treatment phases signifying treatment related improvement
✗ = any change over the study could not be attributed to the treatment
ON= Orthographic Neighbourhood

Discussion: Phase 1.

In Phase 1, Copy and Recall Treatment resulted in improved spelling accuracy for all treated sets in both whole word and letter accuracy analyses. Moreover, this improvement was greater during treated than untreated phases of the study, confirming treatment-related effects. None of the unrelated control sets improved which further reinforces that any improvement on other sets was due to treatment and not due to the effects of practice, placebo or spontaneous recovery. There was no generalisation to any untreated partners in any analysis except for the direct neighbours. Nonetheless, for this set, significant improvement was only observed for the whole-word analysis, suggesting that the effect was not robust and it would be unwise to make inferences on the basis of this data. It became apparent, during the spoken baselines and post-tests that, despite being asked not to, CWS cued his spelling with spoken naming (he always said under his breath the phonemes before spelling). Using a sublexical strategy to spell during treatment could have resulted in improved phoneme-to-grapheme conversions. However, as treatment related improvement was only found for treated items this suggests treatment did affected lexical access (not phoneme-to-grapheme conversion as this would have resulted in generalisation).

At face value, the results of this study suggest that homophone representations are not shared in the orthographic output lexicon, nor is there any (effect of) feedback from graphemes to the lexicon as there was no clear generalisation to untreated items. This is in contrast to previous (spoken) homophone (Biedermann & Nickels, 2008a, 2008b; Biran et al., 2013) and direct neighbour generalisation (Harris et al., 2012; Sage & Ellis, 2006), although in line with Krajenbrink et al. (2016). Given this inconsistency, it is important to consider alternative explanations. One possibility is that perhaps our treatment was not of sufficient duration to maximally improve the treated items (and result in generalisation). Other studies that have found generalisation have used considerably longer periods of treatment and found up to 100% accuracy (80% improvement) on treated items (e.g., Raymer, Cudworth, & Haley, 2003). Some researchers continue treatment until a pre-specified level of performance has been reached , or continue treatment of predefined items until generalisation occurs (Thompson, 1989). Regardless, it is clear that the treatment effects were not long lasting (see Figures 3-5). Once again it is possible that this was due to the duration of the treatment. Howard, Patterson, Franklin, Orchard-Lisle and Morton (1985) note that studies with more intensive and prolonged therapy report more beneficial effects.

Indeed, a longer treatment duration previously resulted in longer lasting (6 weeks post intervention) treatment gains with CWS (Roberts, 2013).

In Phase 1 a large number of items (nearly 100) in each session were treated, using a relatively shallow method (copying). Although research has found that treatment with many items can be just as effective as treatment with fewer items (Snell, Sage, & Lambon Ralph, 2010), it remains possible that the large set of items used in treatment might have interfered with learning, reducing the possible effect of treatment. For example, recall with larger lists of items has resulted in higher intrusion rates (Toglia, 1999). In addition, increased stress has been found to negatively affect the outcome of aphasia therapy (Murray & Ray, 2001). Conceivably then, the perceived pressure of large numbers of items could have negatively affected the results of treatment for CWS. Accordingly, we felt it was worth considering whether treatment would have produced better outcomes if it were to focus intensively on fewer items.

Finally, it seemed that CWS would often be sounding out the correct spelling but writing one letter that was a phonetic and/or an orthographic substitute (e.g. sounding out 'b' 'a' 't' but writing *pat*. It is possible that CWS' other deficits (his mixed dysgraphia profile and dyslexia) could be masking effects of treatment. Perhaps a treatment which would reinforce selection of the correct grapheme over a orthographically/phonologically similar foil would result in maximum treatment effects.

The combination of all these factors led us to conduct a second phase of treatment six weeks after the last post-test of Phase 1. This second phase included slight modifications aiming to address some of the issues raised above: during assessment, we provided both the picture and dictated the picture name; we undertook a more intensive treatment with fewer items, but more sessions, and used a treatment that aimed to invoke deeper processing (anagram with copy and recall therapy) which included selecting correct letters over foils. With these modifications, we aimed to investigate whether we would replicate the pattern seen in Phase 1 or whether Phase 2 treatment would generalise to items with differing degrees of overlap.

Phase 2

Method: Phase 2.

Stimuli.

As before, the treated and untreated subsets were matched for accuracy across baselines and also matched for the various psycholinguistic variables shown in Table 6. As in Phase 1, homophonic and non-homophonic controls with direct neighbour items (subsets 1-6) were randomly assigned to one of two sets. Items were then swapped between subsets to ensure matching on psycholinguistic variables and baseline accuracy (as for Phase 1) but also to ensure similar numbers of items that were treated in Phase 1 appeared in each subset. Once subsets 1-6 were matched across their two sets of partners, they were randomly allocated to be treated or untreated. Subset 7 and 8 were never treated as these were the orthographic neighbourhood controls. All groups were matched to their respective subsets on the variables apart from Subset 8. As in Phase 1, subset 8 was not matched to Subset 5 (and therefore 6 and 7) on orthographic neighbourhood (or any variables confounded with this) as it was designed to have a very small orthographic neighbourhood density.

Procedure.

Phase 2 followed a similar procedure to Phase 1 except that there were eight treatment sessions before and after the interim test (16 sessions in total, rather than 12 in Phase 1). CWS also completed 16 sessions of copy and recall homework during Phase 2 (rather than six in Phase 1). During this homework CWS copied the written name presented beneath the picture, then after an intervening item was asked to recall the picture spelling. Due to CWS's hospitalisation, only one post-test was collected.

It became obvious in Phase 1 that CWS was dictating the picture name prior to spelling. It also became apparent that this did not cause any treatment effects due to sublexical strategies. Therefore, some changes were made to reduce the effects of CWS' anomia and for ease of administration. During the Phase 2 assessment sessions the picture was dictated for writing in the presence of the picture, with the description presented underneath, but not read aloud by the experimenter. As before, baselines were split into two sessions with presentation order counterbalanced, however CWS was no longer tested on his spoken production of the homophones (i.e., baselines and post-test only assessed spelling).

Table 6. The experimental subsets matched on percentage accuracy (acc), log10 frequency (frequency) and other psycholinguistic variables for Phase 2.

	Homographs		Heterograph		Non-homophonic controls			
Subset	1	2	3	4	5	6	7	8
	Treated (N=30)	Untreated partners (N=30)	Treated (N= 13)	Untreated partners (N=13)	Treated (N=25)	Direct neighbours of 5 (N=25)	High Orth N (N=25)	Low Orth N (N=22)
Written accuracy: B1 (%)	30.00	26.67	7.69	7.69	25.00	25.00	32.00	13.60
Written accuracy: B2 (%)	13.33	13.33	15.38	23.08	33.33	33.33	36.00	9.09
Written accuracy: B2 (%)	13.33	20.00	15.38	23.08	29.20	29.17	40.00	4.55
Treated in Phase 1	10	12	8	5	11	13	0	0
Written Lemma	2.15	2.12	2.54	2.55	2.72	2.70	2.86	2.69
Written word form	1.97	1.93	2.41	2.38	2.55	2.52	2.73	2.49
Spoken lemma	0.79	0.72	1.25	1.25	1.18	1.13	1.25	1.01
Spoken word form	0.67	0.59	1.11	1.17	1.03	0.98	1.14	0.84
Syllables	1.17	1.17	1.08	1.08	1.00	1.00	1.00	1.00
Phonemes	3.53	3.53	3.00	3.00	3.21	3.13	3.40	3.18
Letters	4.27	4.27	4.46	4.15	4.04	3.96	3.84	5.05*
Orthographic neighbourhood density	8.53	8.53	7.38	8.54	10.79	10.58	8.72	0.59*
Orthographic neighbourhood freq.	114.54	114.54	64.13	413.03	108.14	99.26	321.63	26.94*
Phonological neighbourhood density	17.07	17.07	20.54	20.17	20.58	20.79	15.08	11.41*
Phonological neighbourhood freq.	106.97	106.97	451.62	478.02	235.14	186.08	274.21*	459.21
Regularity	0.90	0.90	0.50	0.54	0.92	0.92	0.80	0.54*

*Significantly different from matched subset
Orthographic neighbourhood density = number of words with one letter difference
Orthographic neighbourhood freq. = average frequency of all the orthographic neighbours
Phonological neighbourhood density = number of words with one phoneme difference
Phonological neighbourhood freq.= average frequency of all the phonological neighbours

Treatment.

In treatment session 1, 17 items were randomly selected from the each of three subsets of 67 treated items to be 'intensively treated' using a task which was adapted from Beeson's (1999) anagram with copy treatment (ACT). See Figure 6 for a schematic representation of this ACT protocol. CWS arranged the target letter tiles into the correct order before copying. After this, two distractor tiles (letters not present in the target) were presented along with the target letters. Again, CWS was asked to sort these into the target word before copying. After a delay (ten seconds) he was asked to recall the written name of the picture. Session 1 involved ACT therapy of 17 items, followed by copying of all 67 treated items (the 50 non-intensively treated items of that session plus the intensively treated items). This was done in attempt to increase consolidation of the intensively treated items as well as retain improvement on items intensively treated in a previous session (this would obviously only be the case from session 2). A new group of 17 items were intensively treated in each of for treatment sessions 2, 3 and 4 (16 in session 4). Then ACT treatment sessions 5-8 were a repetition of sessions 1-4. This meant that prior to the interim test each item received ACT twice and was copied eight times. Each of the intensively treated sets had roughly equal amounts from each subset (homographs: 7-8, heterographs: 3-4, non-homophones: 5-7).

After the interim test, the treatment was adapted again to include an immediate attempt to spell the name before continuing with the same procedure in Figure 6. This was done to save time and energy as well as promote writing to dictation in the presence of a picture. If the item was spelled correctly it was covered and probed again after a delay before moving on to the next item. If an item was spelled incorrectly the same procedure was adopted as in Figure 6. As before, after 17 items were treated in this intensive section of the session, all 67 treatment items were presented for probing, including the 17 intensively treated items. If any of these items were incorrect they were presented for copying.

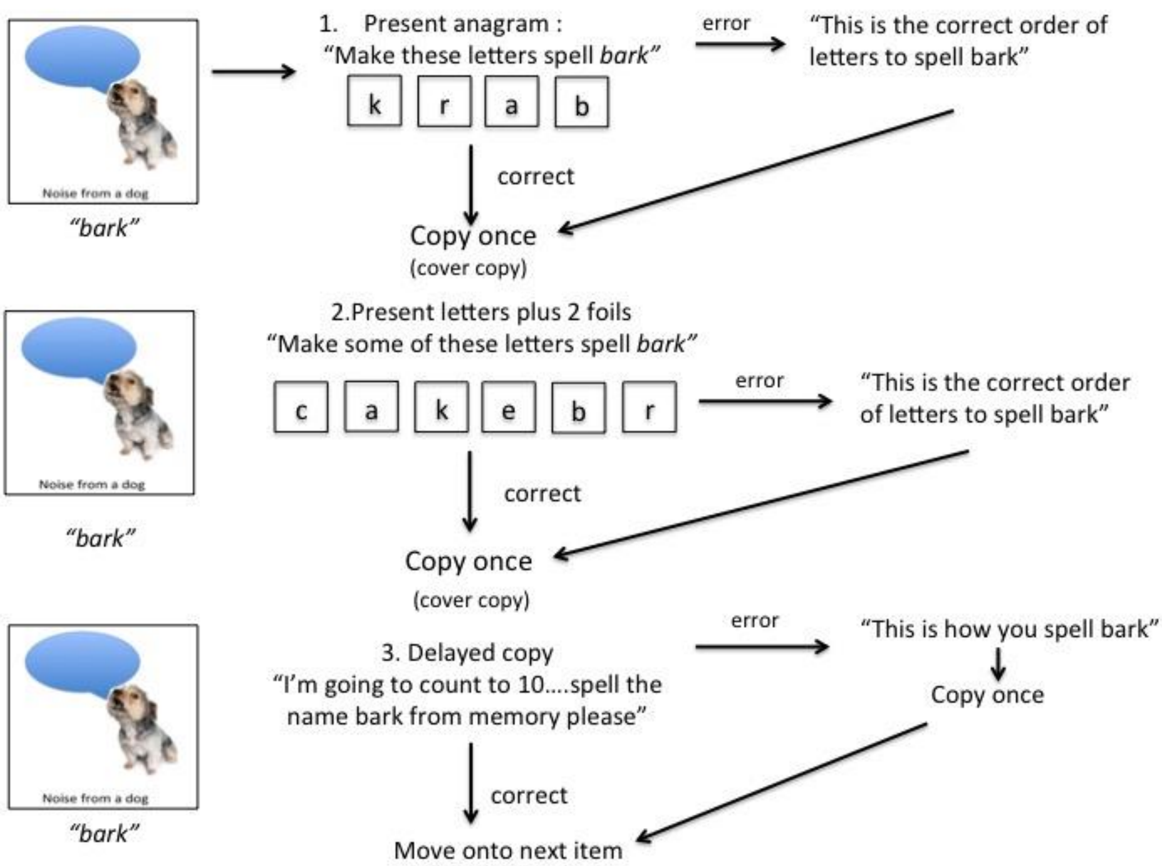


Figure 6: Schemantic representation of the ACT treatment conducted on 17 items during treatment sessions one to eight in Phase 2. Following this procedure all 67 items were presented for copying, including the 17 previously ACT treated items.

Results: Phase 2

The results were scored and analysed in the same way as Phase 1 (accuracy and letter scoring, using both WEST-Trend and WEST-ROC) and are presented in Table 7. As Table 7 and Figure 7 show, like Phase 1 the treatment significantly improved spelling of the treated items in both whole word and letter analysis. However, unlike Phase 1, there was no significant improvement due to treatment for either word or letter accuracy for the untreated partners. Like Phase 1, there was no treatment-related improvement for the unrelated control words (high and low in orthographic neighbours) for either analysis. As there was a clear effect of treatment overall, we therefore investigated treatment-related improvement within the subsets.

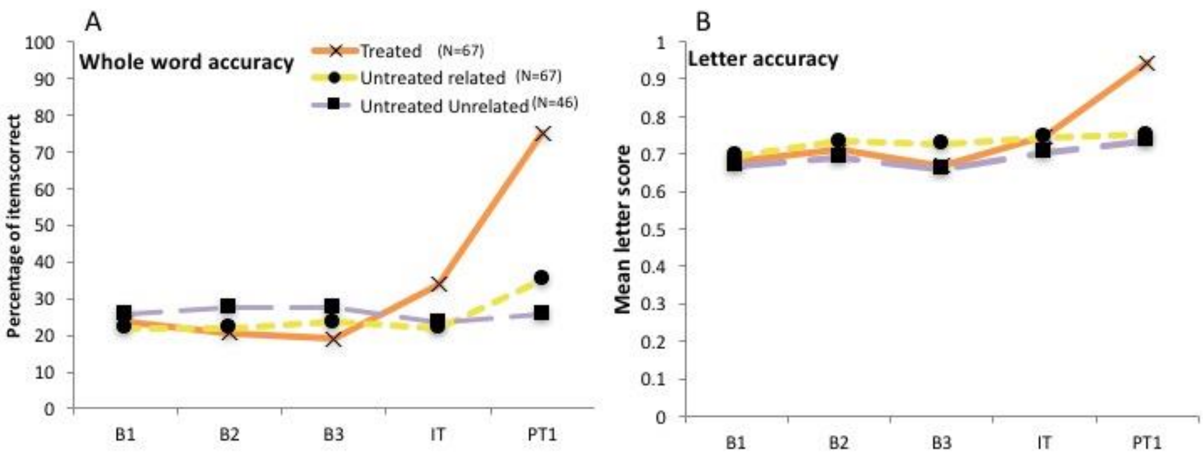


Figure 7: Phase 2 performance scored by a) whole word accuracy, and b) letter accuracy overall letter scoring for the combined treated items (homographs, heterographs and non-homophonic controls), untreated related (homograph and heterograph partners and direct neighbours of non-homophonic controls) and untreated unrelated (high and Low orthographic neighbourhood controls)

Replicating Phase 1, the treated homographs and heterographs improved following treatment but there was no improvement for the untreated partners (see Figure 8).

The effect of treatment on the non-homophonic controls was not consistent across analyses (Figure 9, Table 7). Only the whole word analysis showed significant treatment-related improvement for the treated items. Letter accuracy was close to ceiling and did not show improvement. The untreated direct neighbour partners did not improve in either analysis. The non-homophonic high and low orthographic neighbourhood showed no improvement due to treatment (Figure 9 and Table 7), which was consistent with Phase 1.

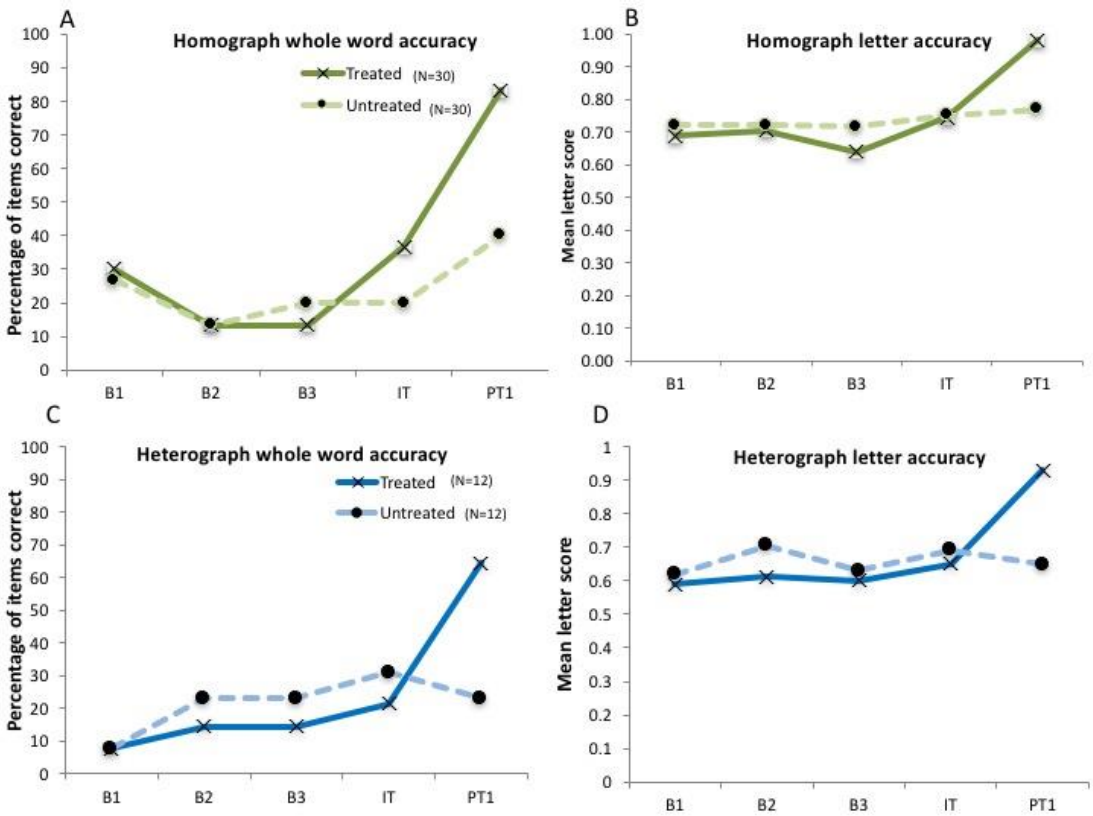


Figure 8: Phase 2 accuracy for homograph (top row) and heterograph (bottom row) on whole word accuracy (A&B), and letter accuracy (B&D).

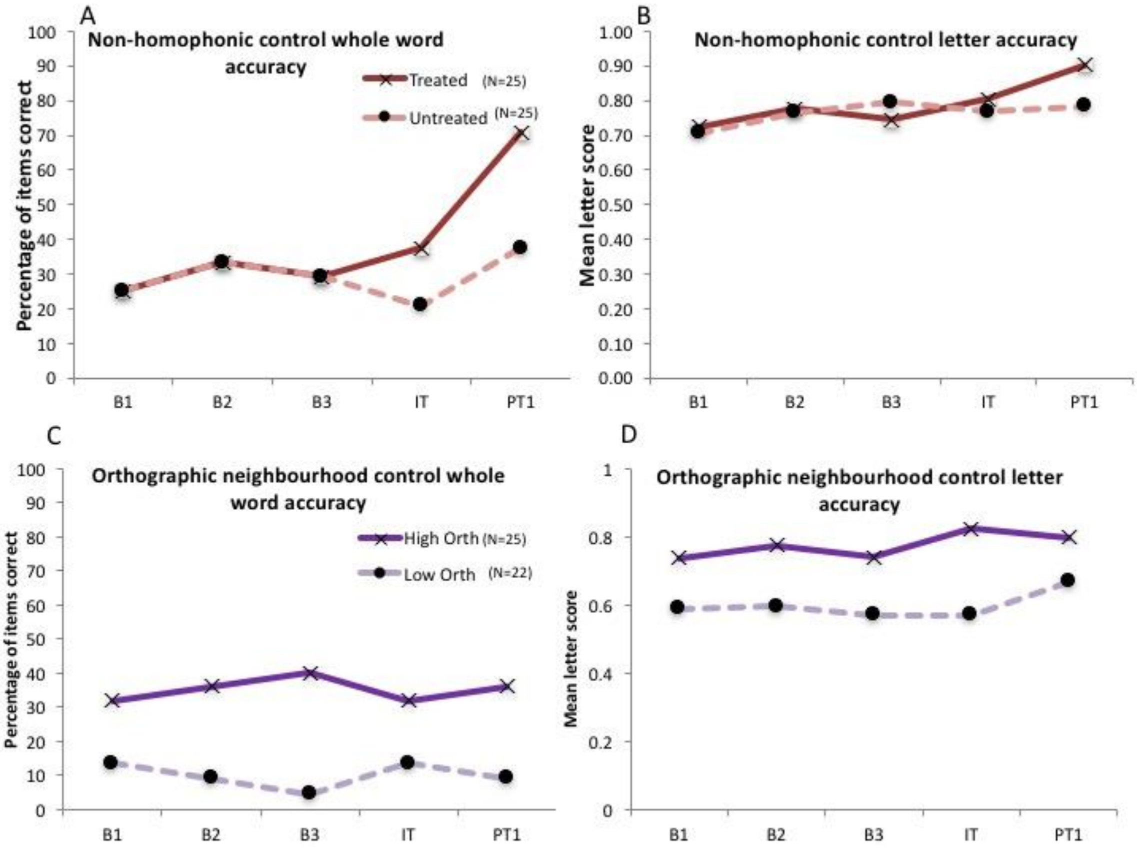


Figure 9: Phase 2 accuracy for non-homophonic controls and direct neighbours (top row) and orthographic neighbourhood controls (bottom row) on whole word accuracy (A&B) , and letter accuracy (B&D)

Table 7. Results from the WEST t-tests and their significance for Phase 2 of the treatment compared to Phase 1 (with one post-test)

	Phase 2 (one post test)							
	Whole word			Letter analysis			Consistency (same findings across ....)	
Set (degrees of freedom)	West trend	West-ROC	Treatment specific improvement	West trend	West-ROC	Treatment specific improvement	Analysis (word and letter accuracy)	Phases (Phase 1 and Phase 2)
Treated (1,59)	8.36**	6.88**	✓	7.96 **	5.42 **	✓	yes	yes
Untreated related (1,59)	1.94*	1.27	✗	2.17*	-0.38	✗	yes	no
Untreated unrelated (1,47)	0.00		✗	2.09*	1.58	✗	yes	yes
Subsets								
Homographs:								
1.Treated (1,21)	6.20**	7.86**	✓	5.76**	5.27**	✓	yes	yes
2.Untreated (1,21)	1.33		✗	1.25		✗	yes	yes
Heterographs:								
3.Treated (1,13)	4.98**	1.89*	✓	5.69 **	3.58*	✓	yes	yes
4.Untreated (1,13)	1.33		✗	0.40		✗	yes	yes
Nonhomophone:								
5.Treated (1,23)	3.70**	2.61*	✗	3.24**	1.31	✗	no	no
6.Direct N (1,23)	0.72		✓	2.29*	-1.27	✗	yes	no
Unrelated								
7. High Orth N (1,24)	0.21		✗	1.94*	0.40	✗	yes	yes
8. Low Orth N (1,21)	-0.20		✗	1.09		✗	yes	yes

*$p<.05$ **$p<.001$
✓= both West-trend and West-roc were significant indicating significant improvement over the course of the study and greater improvement during the treatment than no treatment phases signifying treatment related improvement
✗= any change over the study could not be attributed to the treatment

Discussion: Phase 2.

To summarise, in Phase 2, all subsets showed the same direction of (or lack of) treatment-related change as in from Phase 1. All the treated subsets improved due to treatment for both whole word and letter accuracy, apart from the treated non-homophonic controls, which showed treatment-related improvement only in the whole word accuracy analysis. There was no evidence for any untreated set improving as a result of treatment.

We speculated that perhaps the lack of generalisation in Phase 1 was due to insufficient treatment, the treatment technique or the assessment measures used in that phase. However, when comparing the effects of treatment on untreated related sets from Phase 1 and Phase 2 there was no difference in the improvement trend across the phases for either whole word (WEST-trend, two sample t-tests; t (1,125)=1.32, p=.19) or letter accuracy (t(1,125)=0.26, p=.79). Moreover, the treatment was no more effective for the treated items in Phase 2 than Phase 1 (WEST-ROC, two-sample t-tests; whole word accuracy t(1,125)=0.43, p=.67) or letter accuracy (t(1,125)=1.05, p=.30). This suggests that the change in the type and amount of treatment and different method of baseline assessment did not affect the effects of treatment and hence were unlikely to be the source of lack of generalisation.

General Discussion

The aim of this study was to investigate whether homophones have shared or separate representations in the orthographic output lexicon by investigating generalisation of treatment to items with differing degrees of orthographic overlap. The treated items in both Phase 1 and Phase 2, and across both whole word and letter analysis improved, however there was no robust generalisation.

First, this lack of generalisation supports the view that, as planned, treatment improved access to whole word representations in the lexicon. If treatment had improved phoneme-to-grapheme correspondence rules, or processing of the graphemic buffer (or grapheme level), generalisation would be predicted (at least in the letter analysis) as treated homophones and treated non-homophones had orthographic overlap with their homophone or direct neighbour partners.

Unlike previous studies in the spoken modality (Biedermann & Nickels, 2008a, 2008b; Biran et al., 2013), despite treated homophone naming improving, we did not find any generalisation to improved naming of homophone partners. At first sight, this supports the assumption of independent written representations for homophones: Theories that invoke shared representations for homophones predict that improved lexical access for one homophone should result in improvements in access of its homophone partner. However, this logic only holds when it is assumed that treatment results in improved processing of a representation at the word form level. Alternative accounts of therapy mechanisms and production repetition priming propose that prior activation can improve access to the lexical representation by strengthening the links to these representations (e.g., Howard, Hickin, Redmond, Clark, & Best, 2006). Howard (2000) claimed that effective therapy activates both semantics and phonology thereby strengthening the link between them. If this were the basis of the improvements due to treatment, even if homophones share a representation (as in the Two-stage model, Levelt et al., 1999), generalisation would not be found. Improvement would only be found for the semantic-lexical link that was treated. Hence, the current data either suggests treatment manifested in the links to lexical representations, or that treatment affects the orthographic word form and this is not shared for homophones.

If therapy does target the word form level, it is nonetheless possible that homophones share a representation in the phonological output lexicon, but have separate representations in the orthographic output lexicon. Although at first it seems inconsistent to have different architectures for phonology and orthography, Best, Herbert, Hickin, Osborne and Howard (2002) point out that this seems more plausible when considering the vast differences between acquiring spoken and written language.

Although not obviously parsimonious as to why homophones perhaps share phonological but not orthographic representation. However, as it currently stands, it is not possible to distinguish whether the differences in homophone generalisation in spoken and written modalities are due to distinct lexical architectures, or differences between the participants. In order to fully test this hypothesis, the same participant should undergo both a phonological and orthographic treatment. If indeed homophones are represented as shared representations in spoken, but independent in written modalities, then generalisation should occur for spoken naming treatment but not for written naming treatment.

One difference between CWS and the participants that undertook the phonological homophone treatment (Biedermann et al., 2002; Biedermann & Nickels, 2008b, 2008a) is that they were monolingual, whereas CWS was a Welsh-English bilingual. Bilinguals are known to have smaller vocabulary sizes within each of their languages compared to monolingual speakers (e.g. Bialystok & Feng, 2009). It is conceivable that a late bilingual might not know both homophone meanings, and therefore homophone representation may be different to that of a monolingual who has a larger vocabulary and is familiar with both word forms. However, this seems unlikely for CWS who, while he grew up in a Welsh dominant household, was exposed to English in the community frequently and from an early age. Nonetheless, it would be worthwhile investigating both orthographic and phonological homophone treatment within a monolingual participant to rule out that the possible lack of generalisation was not caused by some undetected non-native-like differences in the English lexicon due to CWS being bilingual[8].

We also found little evidence for generalisation to orthographically related items in general. While there was some significant improvement in untreated direct orthographic neighbours in Phase 1, this was only evident in the whole word analysis, and was not replicated in Phase 2. Sage and Ellis (2006) suggested that generalisation to direct orthographic neighbours originated from treatment increasing activation of the treated word form (e.g., *bath*) activating the graphemes (e.g., *B, A, T, H*). If most of these are shared with a neighbour (e.g., *A, T, H*), activation is fed back to the neighbouring whole-word representations (e.g., *path*), which may prime these representations and improve subsequent access to these items. However, if this were the cause of generalisation for direct neighbours in the whole word analysis in Phase 1, then we should have also found generalisation to untreated written homographs as they share 100% of graphemes with their treated partners. The lack of replication across analysis type and phases, coupled with the lack of theoretical parsimony (i.e., we would expect generalisation to homographs, as well as, if not more than, direct neighbours), means it is unlikely that partial generalisation to direct neighbours in Phase 1 was more than statistical noise.

[8]Gvion, Biran, Sharabi, and Gil (2015) conducted a phonological homophone treatment with a bilingual participant, however, as this participant suffered from phonological output buffer impairment (not phonological word form impairment) homophone generalisation was not predicted. In fact, no treatment effects at all were found in this individual. Therefore this particular case is uninformative in terms of homophone representations and whether being bilingual can influence homophone representation and generalisation.

In sum, we found no evidence of generalisation and this lack of generalisation is in contrast to other orthographic treatment (Harris et al., 2012; Kohnen et al., 2008; Sage & Ellis, 2006; but see Krajenbrink et al., 2016). This seems to suggest either that there is no feedback from graphemes to the lexical level (as suggested by Caramazza et al., 2001; Levelt et al., 1999), or that if there is feedback (as suggested by e.g. Middleton et al., 2015; Schwartz et al., 2006), there was no influence of this feedback on subsequent performance following treatment.

However, the question remains as to why other studies found generalisation but the present study did not. It was noted above that CWS may suffer from an additional graphemic buffer impairment (Roberts, 2013). Although he has a clear lexical deficit, he also shows length effects and makes spelling errors that are not phonologically plausible (i.e. not regularisation errors) which is consistent with graphemic buffer impairment (e.g. Caramazza, Miceli, Villa, & Romani, 1987).

Could this explain the lack of generalisation? Feasibly, treatment improved (access to) treated impaired lexical representations. This would result in increased activation of the graphemes for these treated items in the graphemic buffer. However, with a damaged graphemic buffer the activation may not be fed back to neighbours containing those graphemes, or the amount fed back may not be enough to cause improved access to related word forms. Indeed other studies have also found a lack of generalisation in individuals with graphemic buffer deficits and attributed it to insufficient feedback (Krajenbrink et al., 2016). Nevertheless, both Harris et al. (2012) and Sage and Ellis (2006) found generalisation to direct orthographic neighbours in individuals with graphemic buffer impairment and Rapp and Kane (2002) attributed generalisation in graphemic buffer impairment to improved overall processing speed of the buffer. If this had been the source of the generalisation for Harris et al. (2012) and Sage and Ellis (2006), then improvement of all untreated items, regardless of their orthographic relationship to the treated items would have been expected. However, Sage and Ellis (2006) only found generalisation to direct neighbours and Harris et al. (2012) only found generalisation to neighbours that shared medial positions. This suggests that it was more likely to be feedback to items with specific orthographic overlap to the treated items that caused generalisation in these participants, not an overall processing improvement at the graphemic buffer level.

Why might CWS and the participants in Krajenbrink et al. (2016) differ from those of Harris et al. (2012) and Sage and Ellis (2006) in terms of generalisation? Although not straightforward, one possibility is that the comorbidity of a lexical deficit and a GB impairment results in a lack of generalisation. Indeed, both CWS and the participants in Krajenbrink et al. (2016) show both lexical and graphemic buffer deficit symptoms. However, although JF (Harris et al., 2012) and BH (Sage & Ellis, 2006) were identified as graphemic buffer impaired, they also show some lexical effects. Harris et al. (2012) attributed this to lexical influences on the (impaired) graphemic buffer for JF, whereas, Sage and Ellis (2006) attributed the lexical effects to impaired semantic-lexicon links in BH. It is clear that CWS has a severly damaged orthographic lexicon due to his very poor irregular spelling (18%) and the fact he does not present with a clear lexical superiority effect (similar regular and nonword spelling). Whereas, JF and BF had (partially) intact lexical access, evidenced by significantly better spelling of real words compared to nonwords (JF 62.3% vs 23.4%; BH 52% vs 26%). Conceivably the reduced level of activation fed back to the lexicon in JF and BF may have, nevertheless, been sufficient to improve access. This could also result in easier access in subsequent production for the untreated neighbours. Whereas, with considerable damage to lexical access, as for CWS, perhaps a greater amount of feedback is required to result in improve lexical access than is possible from a damaged graphemic buffer.

It was found that for CWS treatment effects rapidly declined after treatment stopped. This might indicate that treatment had not improved access to the orthographic lexicon from the semantic system but instead improved access later in the lexical system (e.g. lexicon-graphemic buffer links), which has been argued to be relatively short lived in the (phonological) lexicon due to the properties of the two levels semantic-lexical and lexical-segment). Indeed, repetition has a much shorter lasting effect on word retrieval compared to tasks involving semantics (Howard, Patterson, Franklin, Orchard-Lisle, & Morton, 1985). This is also a potential source of impairment for CWS (in the orthographic lexicon) and could perhaps explain the lack of generalisation in the present study compared to previous studies, as well as lack of lasting effects.

To conclude, this study shows that both copy-and-recall-treatment and anagram with copy-and-recall-treatment, can (at least transiently) improve spelling of treated items in a man with impaired access to the orthographic lexicon and possibly comorbid GB deficits. However, these effects did not generalise to either homophone partners or orthographically related words. This could suggest that homophones have individual orthographic output lexicon representations similar to (Caramazza et al., 2001, Figure 1B) or Interactive activation model (e.g., Dell, Lawler, Harris, & Gordon, 2004; Middleton et al., 2015, Figure 1c). However, the nature of the participant's impairment may have meant that he could not benefit from any feedback, and hence we are currently unable to differentiate between feedforward only and interactive theoretical accounts.

Alternatively, it is possible that homophones could share an orthographic output repersentation as in the Two-stage model (Levelt et al., 1999; Figure 1a), but that treatment effects were due to strengthened lemma access or the links the lemma level rather than operating at the word form level. This would also predict the results we found, but once again would be hard to reconcile across the different studies in the literature.

In conclusion, this study was not able to replicate phonological homophone treatment generalisation in the written modality. The differing results across participants make it hard to draw strong conclusions about lexical architecture and further research is required to determine the extent to which differences across participants contributes to these results.

References

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1996). Celex2. *Linguistic Data Consortium, Philadelphia*.

Beeson, P. M. (1999). Treating acquired writing impairment: strengthening graphemic representations. *Aphasiology*, *13*(9–11), 767–785. https://doi.org/10.1080/026870399401867

Beeson, P. M., Hirsch, F. M., & Rewega, M. A. (2002). Successful single-word writing treatment: Experimental analyses of four cases. *Aphasiology*, *16*(4–6), 473–491. https://doi.org/10.1080/02687030244000167

Behrmann, M. (1987). The rites of righting writing: Homophone remediation in acquired dysgraphia. *Cognitive Neuropsychology*, *4*(3), 365–384. https://doi.org/10.1080/02643298708252044

Best, W., Herbert, R., Hickin, J., Osborne, F., & Howard, D. (2002). Phonological and orthographic facilitation of word-retrieval in aphasia: Immediate and delayed effects. *Aphasiology*, *16*(1–2), 151–168. https://doi.org/10.1080/02687040143000483

Bialystok, E., & Feng, X. (2009). Language proficiency and executive control in proactive interference: Evidence from monolingual and bilingual children and adults. *Brain and Language*, *109*(2–3), 93–100. https://doi.org/10.1016/j.bandl.2008.09.001

Biedermann, B., Blanken, G., & Nickels, L. (2002). The representation of homophones: Evidence from remediation. *Aphasiology*, *16*(10–11), 1115–1136. https://doi.org/10.1080/02687030244000545

Biedermann, B., & Nickels, L. (2008a). Homographic and heterographic homophones in speech production: Does orthography matter? *Cortex*, *44*(6), 683–697. https://doi.org/10.1016/j.cortex.2006.12.001

Biedermann, B., & Nickels, L. (2008b). The representation of homophones: More evidence from the remediation of anomia. *Cortex*, *44*(3), 276–293. https://doi.org/10.1016/j.cortex.2006.07.004

Biran, M., Gvion, A., Sharabi, L., & Gil, M. (2013). The Representation of Homophones in the Phonological Lexicon: Additional Evidence from Hebrew. *Procedia - Social and Behavioral Sciences*, *94*, 86–87. https://doi.org/10.1016/j.sbspro.2013.09.041

Blanken, G. (1989). Wortfindungsstorungen und Verbales Lernen bei Aphasie- Eine Einzelfallstudie. *Neurolinguistik*, 2, 107-126.

Brunsdon, R., Coltheart, M., & Nickels, L. (2005). Treatment of irregular word spelling in developmental surface dysgraphia. *Cognitive Neuropsychology*, *22*(2), 213–251. https://doi.org/10.1080/02643290442000077

Buchwald, A., & Rapp, B. (2009). Distinctions between orthographic long-term memory and working memory. *Cognitive neuropsychology*, *26*(8), 724-751.

Caramazza, A. (1997). How Many Levels of Processing Are There in Lexical Access? *Cognitive Neuropsychology*, *14*(1), 177–208. https://doi.org/10.1080/026432997381664

Caramazza, A., Costa, A., Miozzo, M., & Bi, Y. (2001). The specific-word frequency effect: Implications for the representation of homophones in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(6), 1430–1450. https://doi.org/10.1037/0278-7393.27.6.1430

Caramazza, A., Miceli, G., Villa, G., & Romani, C. (1987). The role of the Graphemic Buffer in spelling: Evidence from a case of acquired dysgraphia. *Cognition*, *26*(1), 59–85. https://doi.org/10.1016/0010-0277(87)90014-X

Davis, C. J. (2005). N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods*, *37*(1), 65–70. https://doi.org/10.3758/BF03206399

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283–321. https://doi.org/10.1037/0033-295X.93.3.283

Dell, G. S., Lawler, E. N., Harris, H. D., & Gordon, J. K. (2004). Models of errors of omission in aphasic naming. *Cognitive Neuropsychology*, *21*(2–4), 125–145. https://doi.org/10.1080/02643290342000320

Druks, J., & Masterson, J. (2000). An action and object naming battery. Hove.

Fry, E. (2004). Phonics: A Large Phoneme - Grapheme Frequency Count Revised. *Journal of Literacy Research*, *36*(1), 85–98. https://doi.org/10.1207/s15548430jlr3601_5

Gvion, A., Biran, M., Sharabi, L., & Gil, M. (2015). *Learning from naming treatment on the phonological representation of homophones*. Presented at the American Speech Hearing Association (ASHA) convention, Colorado, USA.

Harris, L., Olson, A., & Humphreys, G. (2012). Rehabilitation of spelling in a participant with a graphemic buffer impairment: The role of orthographic neighbourhood in remediating the serial position effect. *Neuropsychological Rehabilitation*, *22*(6), 890–919. https://doi.org/10.1080/09602011.2012.709872

Howard, D. (2000). Cognitive neuropsychology and aphasia therapy. In *Acquired Neurogenic Communication Disorders: A Clinical Perspective*. Retrieved from http://www.wiley.com/WileyCDA/WileyTitle/productCd-1861561113.html

Howard, D., Best, W., & Nickels, L. (2015). Optimising the design of intervention studies: critiques and ways forward. *Aphasiology*, *29*(5), 526–562. https://doi.org/10.1080/02687038.2014.985884

Howard, D., Hickin, J., Redmond, T., Clark, P., & Best, W. (2006). Re-Visiting 'Semantic Facilitation' of Word Retrieval for People with Aphasia: Facilitation Yes But Semantic No. *Cortex*, *42*(6), 946–962. https://doi.org/10.1016/S0010-9452(08)70439-8

Howard, D., Patterson, K., Franklin, S., Orchard-Lisle, V., & Morton, J. (1985). Treatment of word retrieval deficits in aphasia. A comparison of two therapy methods. *Brain: A Journal of Neurology*, *108 (Pt 4)*, 817–829.

Jacobs, B. J., & Thompson, C. K. (2000). Cross-Modal Generalization Effects of Training Noncanonical Sentence Comprehension and Production in Agrammatic Aphasia. *Journal of Speech Language and Hearing Research*, *43*(1), 5. https://doi.org/10.1044/jslhr.4301.05

Jescheniak, J. D., & Levelt, M., J. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 824–843. https://doi.org/10.1037/0278-7393.20.4.824

Jescheniak, J. D., Meyer, A. S., & Levelt, W. J. M. (2003). Specific-word frequency is not all that counts in speech production: Comments on Caramazza, Costa, et al. (2001) and new experimental data. *Journal of Experimental Psychology: Learning,* *Memory, and Cognition*, *29*(3), 432–438. https://doi.org/10.1037/0278-7393.29.3.432

Kohnen, S., Nickels, L., Coltheart, M., & Brunsdon, R. (2008). Predicting generalization in the training of irregular-word spelling: treating lexical spelling deficits in a child. *Cognitive Neuropsychology*, *25*(3), 343–375. https://doi.org/10.1080/02643290802003000

Krajenbrink, T., Nickels, L., & Kohnen, S. (2016). No effect of orthographic neighbourhood in treatment with two cases of acquired dysgraphia. *Aphasiology*, *0*(0), 1–34. https://doi.org/10.1080/02687038.2016.1218435

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–38.

Middleton, E. L., Chen, Q., & Verkuilen, J. (2015). Friends and foes in the lexicon: Homophone naming in aphasia. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(1), 77–94. https://doi.org/10.1037/a0037778

Murray, L. L., & Ray, A. H. (2001). A comparison of relaxation training and syntax stimulation for chronic nonfluent aphasia. *Journal of Communication Disorders*, *34*(1–2), 87–113. https://doi.org/10.1016/S0021-9924(00)00043-5

Perry, C., Ziegler, J. C., & Coltheart, M. (2002). A dissociation between orthographic awareness and spelling production. *Applied Psycholinguistics*, *23*(1), 43–73. https://doi.org/10.1017/S0142716402000036

Raymer, A., Cudworth, C., & Haley, M. (2003). Spelling treatment for an individual with dysgraphia: Analysis of generalisation to untrained words. *Aphasiology*, *17*(6–7), 607–624. https://doi.org/10.1080/02687030344000058

Roberts, J. (2013). *Cross-Lingisitic treatment generalisation in Welsh-English Bilingual Aphasia* (Unpublished doctoral dissertation). Bangor University, UK.

Sage, K., & Ellis, A. (2006). Using orthographic neighbours to treat a case of graphemic buffer disorder. *Aphasiology*, *20*(9), 851–870. https://doi.org/10.1080/02687030600738945

Schwartz, M. F., Dell, G. S., Martin, N., Gahl, S., & Sobel, P. (2006). A case-series test of the interactive two-step model of lexical access: Evidence from picture naming. *Journal of Memory and Language*, *54*(2), 228–264. https://doi.org/10.1016/j.jml.2005.10.001

Shelton, J., R., & Weinrich, M. (1997). Further Evidence of a Dissociation between Output Phonological and Orthographic Lexicons: A Case Study. *Cognitive Neuropsychology*, *14*(1), 105–129. https://doi.org/10.1080/026432997381637

Snell, C., Sage, K., & Lamborn Ralph, M. A. (2010). How many words should we provide in anomia therapy? A meta-analysis and a case series study. *Aphasiology*, *24*(9), 1064–1094. https://doi.org/10.1080/02687030903372632

Tainturier, M.-J., Roberts, J., & Leek, E. C. (2011). Do reading processes differ in transparent versus opaque orthographies? A study of acquired dyslexia in Welsh/English bilinguals. *Cognitive Neuropsychology*, *28*(8), 546–563. https://doi.org/10.1080/02643294.2012.698986

Thompson, C. (1989). Generalization research in aphasia. *Clinical Aphasiology*, *18*, 195–222.

Toglia, M. P. (1999). Recall Accuracy and Illusory Memories: When More is Less. *Memory*, *7*(2), 233–256. https://doi.org/10.1080/741944069

Appendices

Appendix A

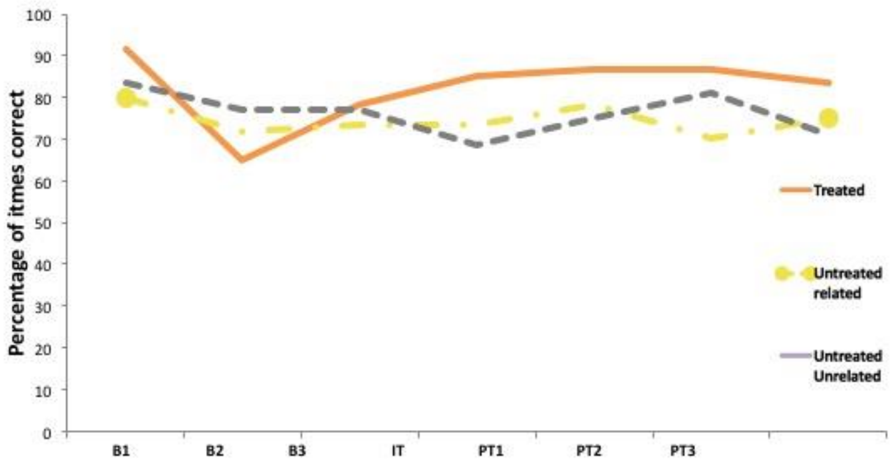


Figure A1: Spoken naming results from Phase 1 for the treated (subsets 1,3 &5), the untreated related (subsets 2,4,6) and the untreated unrelated (subsets 7&8)

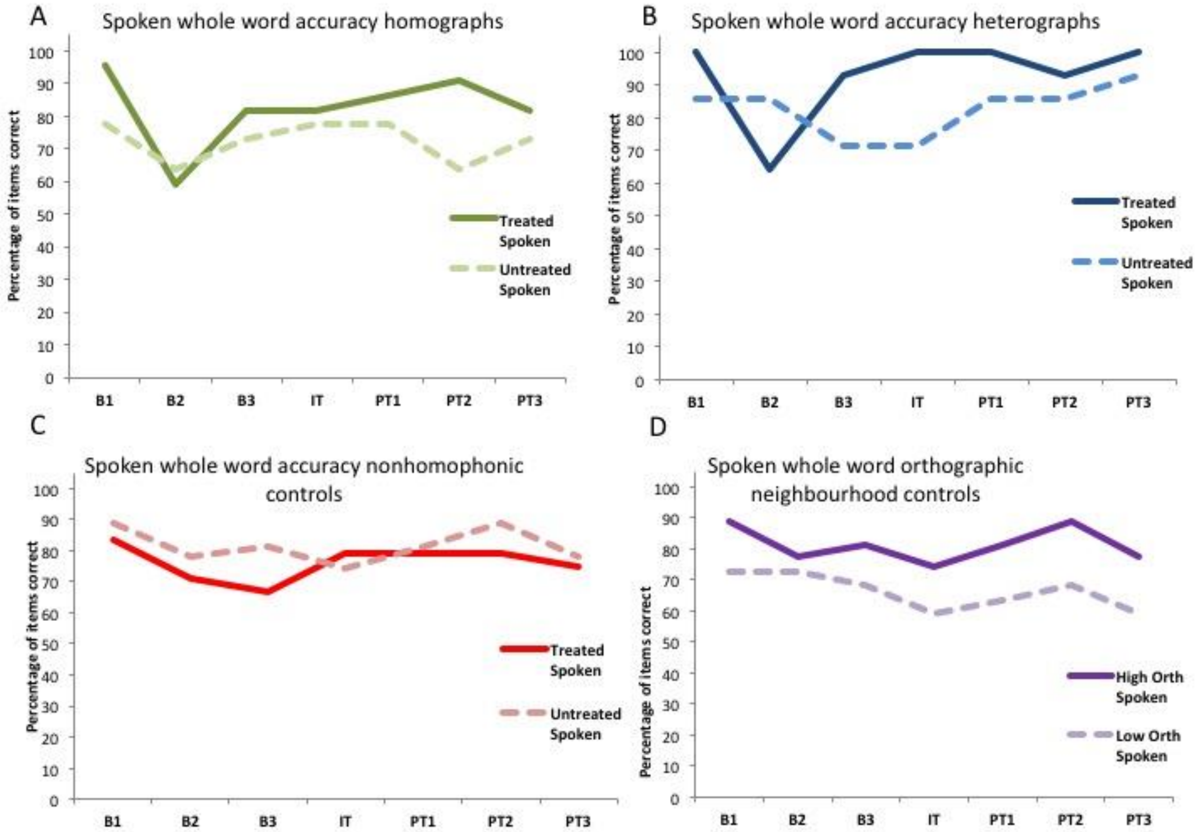


Figure A2. Graphs to show the spoken production of the a) the homographs (subsets 1&2), b) the heterographs (subsets 3 &4) c) non-homophonic controls (5&6) and d) the orthographic controls (7&8)

Table A1. Results from the WEST t-tests (t-values) and their significance for Phase 1 spoken results (one post-test)

	West trend	West-ROC	Treatment specific improvement
Treated (1,59)	0.54		✗
Untreated related (1,59)	-0.12		✗
Untreated unrelated (1,47)	-1.85		✗
Treated Heterographs 1,13)	2.68 *	1.59	✗
Untreated heterographs(1,13)	-0.48		✗
Treated homographs(1,21)	0.25		✗
Untreated Homographs(1,21)	0.59		✗
Treated Direct N(1,23)	-0.90		✗
Untreated Direct N(1,23)	-0.36		✗
High ON(1,25)	-1.11		✗
Low ON(1,22)	-1.58		✗

* $p<.05$ ** $p<.001$

✓= both West-trend and West-roc were significant suggesting significant improvement over the course of the study and during the treatment phase signifying treatment related improvement

✗ = any change over the study could not be attributed to the treatment

Table A2. Results WEST t-tests (t-values) and their significance for Phase 1 of the treatment with three post-tests.

	Phase 1 (one post test)							
	Whole word			Letter analysis				
	West trend	West-ROC	Treatment specific improvement	West trend	West-ROC	Treatment specific improvement	Consistent Type of analysis	Consistent to 1 post-test
Treated (1,59)	5.12**	5.25 **	✓	4.27 **	7.11 **	✓	✓	✓
Untreated related (1,59)	0.95		✗	-0.18		✗	✓	✓
Untreated unrelated (1,47)	0.06		✗	-0.84		✗	✓	✓
Treated homographs(1,21)	3.01 **	3.85 **	✓	1.86 *	3.33 **	✓	✓	✓
Untreated Homographs(1,21)	1.28		✗	0.79		✗	✓	✓
Treated Heterographs 1,13)	2.67 *	3.91 **	✓	1.61		✗	✗	✗
Untreated heterographs(1,13)	-0.55		✗	2.67*	2.35*	✓	✗	✗
Treated Direct N(1,23)	3.61 **	4.72 **	✓	-1.56		✗	✗	✗
Untreated Direct N(1,23)	2.15 *	1.85 *	✓	0.09		✗	✗	✓
High ON(1,25)	0.28		✗	-0.32		✗	✓	✓
Low ON(1,22)	-0.44		✗	-1.06		✗	✓	✓

*$p<.05$ ** $p<.001$

✓= both West-trend and West-roc were significant suggesting significant improvement over the course of the study and during the treatment phase signifying treatment related improvement

✗= any change over the study could not be attributed to the treatment

Table A3. The T-values independent sample 2-tailed t-tests on the effect of treatment (WEST) between each subset in Phase 1 vs Phase 2 (with one post-test)

	Whole word		Letter	
Subset (degrees of freedom	West-Trend	West-ROC	West-Trend	West-ROC
Treated (60, 67)		0.43		1.05
Untreated unrelated (1,66)	1.32		0.26	
UT unrelated (1,46)	0.30		2.90**	1.19
Treated homographs (1,29)	0.83		0.90	
Untreated Homographs (1,29)	0.32		0.09	
Treated Heterographs (1,12)	0.20		0.13	
Untreated heterographs (1,12)	0.57		0.94	
Treated Direct N (1,23)	0.91		1.08	
Untreated Direct N (1,23)	1.26		0.96	
High ON (1,24)	0.63		1.73	
Low ON (1,21)	0.34		2.37*	0.29

*p<.05 **p<.001

Chapter 4

Homophone priming of written picture naming: Witch weigh of righting?

Homophones are unique in lexicalisation: they are words that share a phonological form but refer to two different meanings (such as *bank* [financial institution] and *bank* [side of a river]). Due to this phonological overlap they are interesting in terms of lexical representation. Previous homophone research has been unable to resolve two competing theories about homophone representation. One theory posits that homophones share representations at the word form level (e.g., Levelt, Roelofs, & Meyer, 1999). The opposing theories postulates that homophones have independent representations either at the phonological word form level, (e.g., Caramazza, Costa, Miozzo, & Bi, 2001) or are represented in links between lemmas (lexical-syntactic representations mediating between semantics and phonology) and phonemes (Dell et al., 1997; Middleton et al., 2015). However, the majority of homophone research has investigated the question of homophone representation in spoken production. Comparatively little is known about the representation of homophones in written production, although some authors suggest that they might have similar structures due to similarities in production (e.g., Bonin & Fayol, 2002; Caramazza, 1997). Hence, we will use spoken language models to guide our discussion of written homophone representation.

One way to differentiate between the possible accounts of homophone representation is to investigate priming. Repetition priming is a robust experimental finding across multiple tasks and modalities (e.g., Barry, Hirsh, Johnston, & Williams, 2001; Damian, Dorjee, & Stadthagen-Gonzalez, 2011; Ferrand, Grainger, & Segui, 1994; Wheeldon & Monsell, 1992). In spoken picture naming, facilitatory effects (e.g., faster reaction times) have been found when naming a picture that had been preceded by naming the item (the same picture or a different depiction of the same item or naming to definition) compared to naming of an unrelated item (e.g., Mitchell & Brown, 1988;

Wheeldon & Monsell, 1992). This is true for items presented in the same session (Barry et al., 2001) and for long lags between the prime and target item (e.g., up to 48 weeks, Cave, 1997). Priming has also been found for both spoken and written modalities (Damian et al., 2011).

It is generally agreed that priming of picture naming represents structural change to the language processing system rather than reflecting an episodic memory trace (Monsell, Matthews, & Miller, 1992; Barry et al., 2001). The same pathways and representations that were previously accessed, engage in the same processes. As these processes have been primed production is quicker and less erroneous (Howard, Nickels, Coltheart, & Cole-Virtue, 2006). Repetition priming has been found from definitions to pictures of the same items (e.g., Wheeldon & Monsell, 1992), therefore it cannot be accounted for by perceptual priming of, for example, picture recognition. Similarly, as naming a picture in one language does not prime production in a second language (Monsell, Matthews, & Miller, 1992), priming also cannot be caused by structural changes at the semantic level. Hence, there is a general consensus that the locus of repetition priming is in the lexical system. However, the lexical system consists of more than one component, and the locus of priming with this system is less established (as will be discussed below).

There are two main accounts proposed for the change within the lexical system caused by priming. The first is that priming invokes changes in the threshold of activation required to retrieve a lexical entry (as suggested in the Logogen model of lexical access, Morton, 1969) and the other account proposes changes in weights of connections to representations (e.g., Howard et al., 2006; Oppenheim, Dell, & Schwartz, 2010). In the Logogen model, word retrieval entails activation and selection of a logogen (a lexical entry). The selection of a logogen lowers the activation threshold of this item, so that any subsequent production will require less activation to be produced and so will be facilitated. In the alternative view (e.g., Howard et al., 2006; Oppenheim, Dell, & Schwartz, 2010) a lexical item's retrieval speed is determined, in part, by its connection strength from its semantics to its phonological form. The strength (or weight) of each connection determines the activation needed for retrieval. These weights become stronger as a result of prior use, which causes repetition priming effects. These two accounts interact with the nature of homophone representation to make different predictions for homophone priming as discussed below.

Spoken homophone priming.

As noted above, Wheeldon and Monsell (1992) found that spoken naming of a picture after previous spoken production of the same item (either to a written definition or reading aloud) results in faster production time when later naming a picture of the same (non-homophone) target word (repetition priming). Wheeldon and Monsell (1992) discuss how repetition priming is relatively long lasting compared to form priming (e.g., priming with phonological or orthographic form overlap). Moreover, Wheeldon (2003) demonstrated an *inhibitory* effect of form priming, but finds that this is no longer evident with even one intervening item.

Based on the assumption that homophones share a representation at the word form level, Wheeldon and Monsell (1992) hypothesised that, if priming affects word forms (rather than lexical links), homophones should be primed by their homophone partner, just as in repetition priming. Across two experiments, there was no evidence for homophone priming for heterographs. However, the data was not entirely clear. For example, in Experiment 3, there was an interaction between mean reaction times (RTs), prime type (repetition or heterograph) and priming (primed or unprimed) but this interaction was not found when analysing medians. This indicates that, while there was significant priming for repetition primes but not for heterographs (i.e., lack of facilitation when naming picture of *knight*, after prior production of *night* to a definition), the difference between the extent of priming in each condition was not reliable. For homographic homophones, there was a significant facilitatory priming effect. For example, naming a picture of *bank* (the institution) after naming a definition with *bank* (the side of a river) produced faster RTs than when preceded by naming an unrelated item. Once again, however, this was only reliable in the analysis of mean RTs: no significant homograph priming was found when analysing medians.

While these provide some suggestion that facilitatory homophone priming may occur in spoken production dependent on homophone spelling, the results are weak at best (there was also no significant three-way interaction between prime type, homophone type and priming). However, subsequent studies *have* found evidence for homophone priming.

In a similar priming task, older and younger participants produced a homophone (e.g., (cherry) *pit*) in response to a definition, which in turn facilitated proper noun homophone picture naming (e.g., (Brad) *Pitt*; Burke, Locantore, Austin, & Chae, 2004). Both accuracy and reaction times were facilitated in younger participants, whereas only accuracy was affected positively in older participants.

The masked priming paradigm (Forster & Davis, 1984) involves initial presentation of a set of meaningless symbols (e.g., #####) before the target stimuli is presented. However, sandwiched in-between these symbols (the mask) and the target, a prime is presented so quickly that it is consciously undetectable (e.g., Bar & Biederman, 1998). In a French masked priming paradigm, Ferrand, Grainger and Segui (1994) found that timed picture naming (and word reading) of the target (e.g., *pied*) were facilitated by a masked pseudohomophone prime (e.g., *piez*) compared to non-words with orthographic overlap (e.g., *pien*) and unrelated words that shared the initial phoneme (e.g., *peul*). This priming effect from nonword primes was hypothesised to occur due to feedback of activation from sublexical units to whole word representations (Ferrand et al., 1994).

The above studies show that activation of one homophone may produce facilitatory effects on subsequent production of the other homophone.

Phonological aphasia treatment (that aims at improving access to phonological word forms) is analogous to priming of phonological word forms and can facilitate subsequent production. Previous aphasia treatment studies suggests that prior production of one homophone improves access to the trained homophone (item-specific effect) and this generalises to the other homophone partner (e.g., Biedermann & Nickels, 2008a, 2008b). The authors concluded this was due to improved access to a shared word form representation. However, in contrast to Wheeldon and Monsell (1992) this requires localisation of priming (or treatment) effects to changes in the levels in activation of the word form rather than changes in connection strength in the links from semantics to the phonological form.

In sum, homophone priming effects in spoken word production appear to be inconsistent and further investigation in a different modality is warranted.

Written homophone priming.

Like Wheeldon and Monsell (1992), Damian et al. (2011) found that written picture naming is also primed by previously writing the same word to definition. They also replicated the same effect in spoken naming, suggesting that comparable priming processes occur across lexicons. Martin and Barry (2012) discovered that non-word spelling to dictation is influenced by the spelling of a previous word (i.e., when /*vi:m*/was preceded by *theme*, the target was more likely to be spelt as *veme*, whereas when *dream* was the prime, spelling was more likely to be *veam*). The authors concluded this was partly due to feedback from graphemes to the lexicon. Feedback from graphemes to orthographic lexical entries influencing subsequent production is also supported by neuropsychological literature showing generalisation of spelling treatment to direct neighbours (e.g., clock-click; Harris, Olson, & Humphreys, 2012; Sage & Ellis, 2006).

To our knowledge, there has been little attention to written homophone priming in the literature. Jacoby and Witherspoon (1982) found that when both unimpaired and participants with Korsakoff amnesia produced a low frequency heterographic homophone orally in response to a question, they were more likely to produce the primed homophone spelling in a subsequent dictation task than the (target) high frequency homophone spelling. The increased likelihood of the previously produced heterograph being chosen for subsequent production was hypothesised to occur through the initial production of the (prime) homophone making this representation more accessible subsequently when presented with a phonologically ambiguous stimulus. This implies that spoken production (naming) of one member of a homophone pair may prime its corresponding orthographic form, perhaps due to bilateral links between the two lexicons or through simultaneous semantic to orthographic activation (along with semantic - phonological activation) in response to a question. However as this experiment did not include an unprimed comparison condition, conclusions should be taken cautiously. Nonetheless, this account is supported by a study by White, Abrams, Zoller, & Gibson (2008) who found increased homophone spelling errors in writing sentences to dictation (e.g., writing *beech* as *beach*) when the homophone partner had been primed by an orthographically and/or phonologically related prime (e.g., *teacher*).

Finally, however, our study that looked at generalisation from orthographic treatment of homophones in a participant with dysgraphia (Chapter2) did not replicate spoken homophone generalisation described earlier, and again suggests complexity in the results for homophone priming in the written modality, as for the spoken modality.

The present study

Given the inconsistent evidence from the literature to date on homophone priming and the paucity of written priming data in particular, this study aimed to complement the dysgraphia treatment study presented in Chapter 3, by using priming (rather than treatment) to investigate homophone representation. It aimed to address two questions that remain unanswered:

i) Does homophone priming occur in written production?

ii) What is the mechanism underlying homophone priming (if any)?

We will address this issue by comparing homophone primes with repetition (identity) primes and orthographically related (neighbour) primes. Table 1 summarises the predictions from the three main theoretical accounts under different priming assumptions. First, if homophones share a representation at the orthographic word-form level (similar to Levelt et al.'s, 1999, proposal for spoken word production), and priming causes changes in accessibility of word form representations, then we expect priming from homophones to be comparable to that from repetition, but greater than from direct neighbours, which should be equal to unrelated primes. If priming is however, caused by changes of activation in the links from semantics to independent orthographic representations, then we would not expect homophone priming.

Table 1. The predictions of homophone priming compared to other types of priming depending on the level of priming for each possible representation of homophones

	In the semantic-lexical links				**In the representations**			
	Homophone priming	Repetition	Direct neighbour	Unrelated	Homophone priming	Repetition	Direct neighbour	Unrelated
Shared		↑	=	=	✓	=	↓	↓
Separate (no feedback)		↑	=	=		↑	=	=
Separate (feedback)		↑	=	=	✓	? (↑)	? (↓)	↓

✓homophone priming
no homophone priming
↑ more priming than homophones
= equivalent amount of priming to homophones
↓ less priming than homophones

If homophones have independent word form representations in a strictly feedforward model (e.g., the Independent Network model, Caramazza, Costa, Miozzo, & Bi, 2001), then no homophone priming is expected, regardless of whether priming manifests at the word form level or in the semantic-word form links.

The final proposal consists of homophones that have independent lemma representations with orthography represented in the links between these lemmas and grapheme nodes (e.g., Dell et al., 1997; Middleton et al., 2015) with interactive activation between levels. This would seem to predict that if priming is located in the semantic-lemma links, then, like the other models, we expect only repetition priming. However, if priming is represented at the level of orthographic representation (i.e., the lemma-grapheme links), there should be facilitation from homophone primes as the target inherits the activation from the primed partner via spreading activation and feedback. Moreover, this is the only model that also predicts priming of orthographic neighbours through the same mechanism. The level of priming will be directly related to the degree of orthographic overlap (i.e. more priming for homographs than heterographs that should be more similar to orthographic neighbours). However, while the orthographically related target should receive more activation from the prime (compared to when it is unprimed), given that the prime may also be a more effective competitor to the target (if there is competition in the model as postulated by Middleton et al., 2015), the precise pattern of priming is difficult to predict in this model. There also will perhaps be a greater repetition priming effect from homographs due to dissipating activation travelling through the links.

Finally, in terms of homophone spelling: if heterographs are represented in the same way as homographs, we expect their pattern of priming to be the same except perhaps in the case of an interactive activation account, where reduced priming for heterographs should occur compared to homographs due to fewer shared graphemes. Comparing heterographs and homographs will inform us if homophone priming (if any) stems from shared representations or feedback.

In order to answer these questions and test these predictions, we will compare repetition, homograph, heterograph, direct orthographic neighbour (words with one grapheme different e.g., cave-cake) and unrelated priming of written picture naming in unimpaired adults.

Method

Participants

Participants were recruited from Bangor University and received course credit for participation. They were all monolingual English speakers who reported to have grown up in a monolingual household, and only moved to North Wales (from the rest of the UK) for higher education (i.e. over the age of 18). Twenty-five participants completed the experiment. However, data from only 11 participants are reported here (see below) (mean age= 24.18 (4.42), five females, six males).

Materials

The same stimuli used in Chapter 3 were selected for five experimental groups each with a different prime-target relationship: 1) repetition (identity) primes (*church-church,* 19 pairs); 2) homographic primes (*mole* [animal]- *mole* [skin], 30 pairs); 3) heterographic primes (*flower-flour,* 16 pairs); 4) direct neighbour primes (*path-bath,* 25 pairs) and 5) unrelated primes (*bus-tent,* 20 pairs). Stimuli were presented as photographs (300 x 300 pixels) with a written description underneath. The written descriptions were short and were designed to clarify ambiguous pictures (e.g., 'mop' – used for cleaning). The descriptions were included as this study was designed as a parallel study to the aphasia treatment study in Chapter 3. All stimuli (with definitions) had at least 70% name agreement from ten control participants (mean age 29.20 years). In the repeated condition different exemplars of the same item were used (e.g., there were two different photos of the same item that elected the response church, one for each set). This was to reduce any effects due to perceptual priming; instead any effects we find should be due to prior lexical retrieval of the concept and word form named. Each experimental group was split into two sets, A and B, such that each item in Set A had a partner in Set B. See Appendix A for a list of primes and targets.

Matching

Frequency values were obtained from CELEX and log10 transformations were performed. Syllable and phoneme length, orthographic and phonological neighbourhood and neighbourhood density were obtained from N-watch (Davis, 2005). Each item was categorised as irregular or regular for spelling. An item was categorised as irregular if it had at least one grapheme that was not the most frequent for the phoneme it represented in the list by Perry, Ziegler and Coltheart (2002) or was labelled as rare by

Fry (2004). One item from each pair of items was arbitrarily assigned to Set A or Set B. Partners were then adjusted across sets to ensure that, for each experimental group, Set A was matched to Set B for: log written and spoken word form and lemma frequency, number of syllables, number of phonemes, number of letters, number of orthographic neighbours, number of phonological neighbours, average frequency of these orthographic and phonological neighbours and regularity (data for each set on all of these factors is presented in Appendix B).

Procedure

Participants were instructed to look at the stimuli on the computer screen and to write down the first picture name that came to mind. Participants wrote responses on a piece of paper secured onto a Wacom tablet using a Wacom ink pen (stylus). They were told to only read the description if they did not know the picture name and were instructed to begin each trial with the stylus, just above, but not touching the tablet and to write as quickly as possible. The stimulus remained on the screen throughout the trial until the participant signalled they were ready for the next trial by touching the stylus on a green box to the left of the response box. The next item was then presented after a 1000ms blank screen with a fixation cross in the centre. The experiment was run and extracted using Eye and Pen® (Alamargot, Chesnet, Dansac, & Ros, 2006).

There were four different orders within sets that were counterbalanced across participants. Each participant was tested in two sessions at least seven days apart. Each session contained two blocks in which Block 1 comprised primes (preceded by 4 practice items) and Block 2 comprised targets. Block 2 followed Block 1 immediately after a break and a practice item. Participants were randomly assigned to whether Block 1 contained Set A or Set B. Items in Block 2 were presented in the same order as Block 1, therefore there were the same number of intervening items between each prime and its target partner (n=111 including an practice first item that was excluded from analysis). Session 2 reversed the set presentation order for each participant such that participants for whom set A was in Block 1 in Session 1 had Set B in Block 3 (the first block in Session 2).

Analysis

The data was extracted using Eye and Pen ®(Alamargot et al., 2006). All 25 participants completed the experiment, unfortunately however, due to a combination of software and hardware failure it was only possible to retrieve data from all sessions for eleven of the participants. Hence, the results reported here are from this subgroup (4 of whom named Set A first and 7 Set B first)[9].

We removed any items that were incorrectly spelled, where no response was given, or were unable to be extracted due to software failure (23.14%). An incorrect answer included any self-corrections, hence only the first response was scored. Onset reaction times that were clearly mistriggers (i.e., the onset was under 300ms) were removed. Items from the second presentation (target) were also excluded from analysis if their prime was excluded from the first presentation of a session. Two item pairs (one homograph and one heterograph) were removed from analysis as they only reached 40% accuracy overall.

Table 2 gives an example of each prime and target across the experiment. The critical analysis here is within-item priming (i.e., the difference between the same item unprimed e.g., $bath_{up}$ and primed $bath_{p}$. It is not appropriate to examine inter-item priming (i.e., the difference in reaction times between a prime e.g., $bath_{up}$ and a target

[9] I am aware that this is far from ideal and this small number of participants means that the reliability of the results of the analyses may not be strong. I remain hopeful that we may yet be able to retrieve more of the data, or, alternatively run more subjects. Unfortunately, the experiment must be run in Wales (where the equipment is located, and the dialect of the participants match the stimuli) and I am currently based in Australia. As an international student, I am required to submit before my candidature ends (18/11/2016), hence it was not possible for me to go back to Wales and collect more bilingual data before submission. Consequently, this experiment must be considered pilot work.

e.g., $path_p$). As this is a repeated measures design, investigating within-item priming inevitably introduces the confound of practice effects. It is logical to assume that, over time, the participants will get quicker at the task due to, for example, faster recognition of the picture or task practice effects (i.e., independent of lexical priming effects, Block 3 will be quicker than Block 2). Therefore simply comparing unprimed to primed will not take into account the fact that half of the unprimed items (i.e., the items in Block 3) are actually produced after their matched primed items (i.e., the items in Block 2). In order to control for this, we include 'presentation order' as a covariate (order 1: unprimed (Block 1), primed (Block 4); order 2: unprimed (Block 3), primed (Block 2).

Table 2. Order Of Primes And Targets

	Session 1		**Session 2**	
	Block 1	Block 2	Block 3	Block 4
	Prime	Target	Prime	Target
Repetition	$Church_u$	$Church_p$	$Church_u$	$Church_p$
Homograph	Mole $(freckle)_U$	$Mole(animal)_P$	Mole $(animal)_U$	Mole $(freckle)_U$
Heterograph	$Flower_u$	$Flour_p$	$Flour_u$	$Flower_p$
Direct Neighbour	$Bath_u$	$Path_p$	$Path_{up}$	$Bath_p$
Unrelated	Bus_u	$Tent_p$	$Tent_{up}$	Bus_p

U = Unprimed
P= Primed

The statistical software R (R Core Team, 2015) was used with the packages MASS, *lme4* (Bates et al., 2014), and lsmeans (Lenth, 2016). A boxcox test (Box & Cox, 1964) was run on a linear mixed effects model (lme) which resulted in the raw onsets being transformed by -0.3. Competitive modelling was conducted in a stepwise fashion keeping any variables in the model that were significant and replacing any that were not. This resulted in a final model containing group and prime as fixed factors with participant and item number as random slopes and number of syllables, presentation order and name agreement as covariates (see Appendix C for all the models compared). Planned contrasts investigated the effects of prime effect (unprimed vs primed), group prime effect (e.g., unprimed homographs vs primed homographs) and interactions between different group prime effects on inverted RTs, whilst controlling for orthographic neighbourhood, presentation order and name agreement.

Results

As shown in Table 3, there was a main effect of priming (i.e., overall the primed items were written significantly faster than the unprimed items). Figure 1 and Table 3 show that when investigating priming for each group, the repeated items, homographs and direct neighbours were all written significantly faster in the primed condition. However, there was no significant priming found for items primed by heterographs or unrelated items.

When comparing the priming effects between groups, only the repeated items showed a significantly larger difference between primed and unprimed compared to the unrelated condition. There was no significant difference in the extent of priming between any other experimental group (homograph, heterograph, and direct neighbours).

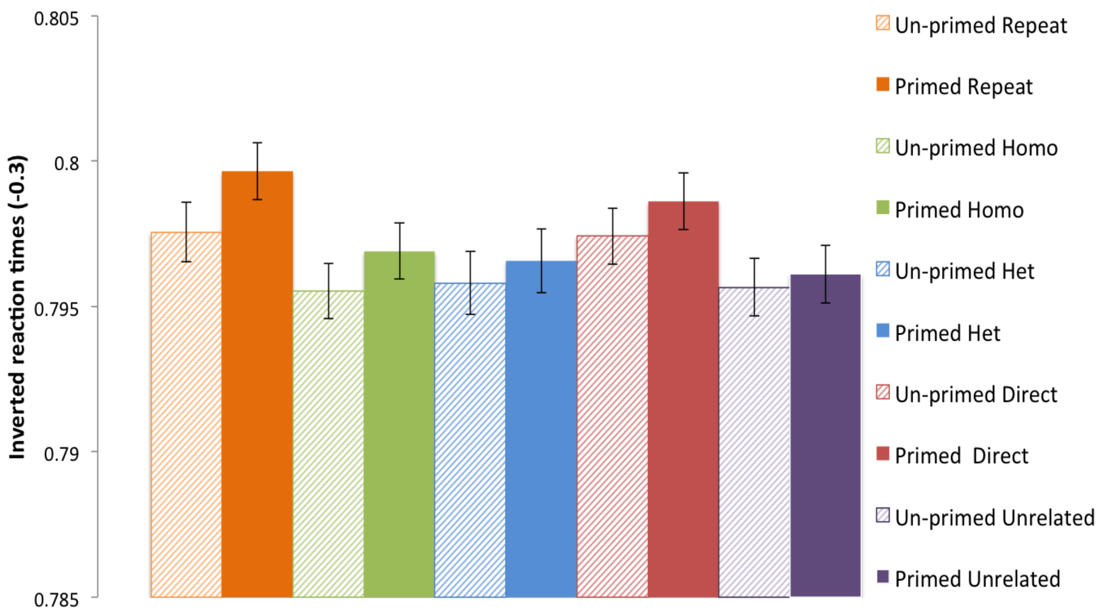


Figure 1. The least squared means for each presentation of each group plotted by session (note that RTs are inverted)

Table 3. Contrasts between the least squared means for the different groups and interactions between presentation and spelling (significant effects in bold).

Analysis					Priming effect
Overall					
Priming (unprimed vs primed)	**$\hat{\beta}$=-5.90^{e-03}**	***SE*=1.01^{e-03}**	***Z*=-5.82**	***p*<.001**	
Repetition	**$\hat{\beta}$=-2.10^{e-03}**	***SE*=4.56^{e-04}**	***Z*=-4.61**	***p*<.001**	
Homograph	**$\hat{\beta}$=-1.36^{e-03}**	***SE*=3.84^{e-04}**	***Z*=-3.55**	***p*<.001**	
Heterograph	$\hat{\beta}$=-7.72^{e-04}	*SE*=5.47^{e-04}	*Z*=-1.41	*p*= .16	
Direct	**$\hat{\beta}$=-1.20^{e-03}**	***SE*=4.05^{e-04}**	***Z*=-2.97**	***p*<.001**	
Unrelated	$\hat{\beta}$=-4.56^{e-04}	*SE*=4.50^{e-04}	*Z*=-1.01	*p*= .31	
Repetition*Unrelated	**$\hat{\beta}$= 8.22^{e-04}**	***SE*=3.20^{e-04}**	***Z*= 2.56**	***p*= .01**	
Homograph*Unrelated	$\hat{\beta}$=-4.54^{e-04}	*SE*=2.96^{e-04}	*Z* =-1.535	*p*= .12	
Heterograph* Unrelated	$\hat{\beta}$=-1.58^{e-04}	*SE*=3.54^{e-04}	*Z*= -0.45	*p*= .66	
Direct *Unrelated	$\hat{\beta}$=-3.73^{e-04}	*SE*=3.03^{e-04}	*Z*= -1.23	*p*= .22	
Repetition*Homograph	$\hat{\beta}$= 3.68^{e-04}	*SE*=2.98^{e-04}	*Z*= 1.24	*p*= .22	
Repetition*Heterograph	$\hat{\beta}$= 6.65^{e-04}	*SE*=3.56^{e-04}	*Z*= 1.87	*p*= .06	
Repetition*Direct	$\hat{\beta}$= 4.49^{e-04}	*SE*=3.05^{e-04}	*Z*= 1.47	*p*= .14	
Heterograph*Homograph	$\hat{\beta}$= 2.96^{e-04}	*SE*=3.34 $^{e-04}$	*Z*= 0.89	*p*= .37	
Heterograph*Direct	$\hat{\beta}$=-2.15^{e-04}	*SE*=3.41^{e-04}	*Z*=-0.63	*p*= .52	
Homograph*Direct	$\hat{\beta}$= 3.31^{e-06}	*SE*=4.88^{e-04}	*Z*= 0.01	*p*=1.00	

= no significant priming effect or interaction
= significant priming effect or interaction

Discussion

This study investigated homophone priming of written picture naming compared to priming from repetition and direct neighbours. The aim was to inform the debate on homophone representation and the level(s) at which priming operates through the pattern of priming across conditions as summarised in Table 1 (earlier). Unfortunately, only a small proportion of the data collected was available for analysis and, hence, this study has to be considered underpowered. Nevertheless, it was clear that prior written picture naming of the same item (repetition condition) resulted in facilitation of subsequent written picture naming over 100 items later. This facilitation was also significantly greater than that found for items primed with unrelated items. This long lasting effect of repetition priming replicates previously found facilitatory effects of repetition in the literature (e.g., Barry et al., 2001; Damian et al., 2011; Wheeldon & Monsell, 1992).

The results were less definitive for the other conditions, while the same form (homograph) or the same form with one letter difference (direct neighbours) resulted in significant facilitation of subsequent written picture naming, these conditions did not show effects that were significantly different to the unrelated condition. Similarly, while prior production of a heterograph did not result in significant priming, this was not significantly different from either the unrelated or the repetition conditions (although there was a trend towards a significant difference compared to the repetition condition). Our homophone priming effects are similar to those found by Wheeldon and Monsell (1992): Both their study and that reported here found weak evidence for homograph priming and no evidence for heterograph priming.

Although we can only report with confidence a repetition priming effect, there is some suggestion of orthographically related priming (from homograph and direct neighbour primes). However, this pattern of priming was not significantly different to other groups (either unrelated or repetition). Nevertheless, we will speculate how this pattern of results might inform location of priming and homophone representation. These discussions are designed to stimulate further research that can test our speculative hypotheses.

Wheeldon and Monsell (1992) concluded that repetition priming was a result of priming semantic-lexical links. Although this could be the case for the repetition priming we found, this cannot explain priming to orthographically related items. The various theories contrasting homophone representation and priming location were presented in Table 1 (earlier) alongside the resulting predictions regarding priming effects. If we assume that our finding of homograph priming will be robust in a larger sample, then this suggests that priming is not solely located in the semantic-lexical links, instead that priming is located at the level of the word form (or links from it).

Homograph priming could be caused by either a shared representation at the word form level (akin to the Two-Stage model, Levelt et al., 1999) or that homographs have independent representations with feedback from shared graphemes (comparable to the Dual Nature account, Middleton et al, 2015). Homophone priming is not possible in a model that contains independent representation with no feedback (i.e. Independent Network model, Caramazza et al., 2001). While priming of a shared representation could explain homograph priming, it does not predict direct neighbour priming. In addition, we would expect to find equivalent priming from homographs and repetition. As this was not the pattern found here, it seems plausible that the homograph priming found was due to feedback from graphemes to independent representations - this could also result in priming for direct neighbours, as was found here.

However, it is hard to predict the precise pattern of priming as a result of treatment and the effects of this priming: when there are independent representations with feedback any orthographically related word could be a competitor at the lexical level. In a competitive interactive model (e.g., Middleton et al., 2015), activation of a target (e.g. bath) would result both in activation and inhibition of potential competitors (e.g. path). The exact effects of the previous activation and inhibition on subsequent production of the competitor are hard to predict as they would depend on a) the exact mechanism of priming (e.g., changes in weights of connections; Howard et al., 2006, Oppenheim, Dell, & Schwartz, 2010 or changes in activation thresholds; Morton, 1969), and b) the precise balance between activation and competition - the previously produced target (e.g. bath) is a stronger competitor as it now has a higher resting state. Hence, in a model that assumes both feedback and competition, any facilitation of previously producing a neighbour (due to spreading activation feeding back resulting in activation and priming) could be counteracted by competition (between similar word forms). Belke, Meyer and Damien (2005) and Schnur, Schwartz, Brecher and Hodgson (2006) discuss in more detail how prior activation of a word can affect subsequent production through competition and inhibition, albeit in terms of semantic retrieval in spoken production, not word form retrieval during written production. The balance between facilitation and inhibition from a prime is discussed briefly in Vitevitch (2002), and Meyer and Bock (1992) discuss the mechanism behind (phonologically) similar word forms competing for selection. Mirman, Dell and Kittredge (2011) explain how many distant phonological neighbours can facilitate production but competition from one close neighbour (a homophone) at the word form level can result in hindrance. However, Gordon (2002) found that phonological neighbours can facilitate production in aphasic speakers (and attributed this as evidence of the Interactive Activation model e.g. Dell, Schwartz, Martin, Saffran, & Gagnon, 1997). In summary, it is clear there is some evidence that competition between word forms can potentially affect production in the phonological lexicon, but also evidence than phonological neighbours can facilitate production. Even less is understood about the role of orthographic neighbours in written production.

With the current data set lacking in power it is not possible to untangle if the lack of significant priming in orthographically related items is due competition masking feedback (or if there is feedback at all). A larger dataset looking at the difference between the controls with a high neighbourhood density and low neighbourhood density as well as the effect of many distant compared to few close (e.g. Mirman et al., 2011) may be able to investigate the level of competition between previously produced items. If competitive interference was substantial we expect there to be no benefit from high neighbourhood density or one close neighbour (i.e. homograph or direct neighbour) as any priming effect would be diminished by activation feeding back and resulting in interference from neighbours. This would also add to the literature on competition within the Interactive Activation model (e.g. Dell, Schwartz, Martin, Saffran, & Gagnon, 1997). However, if there was priming with minimal competition, we might expect a benefit from having lots of orthographic neighbours compared to having very few or one close neighbour (homograph/direct neighbour).

Nevertheless, our data is consistent with feedback within the lexicon from graphemes to word forms, which is supported by numerous findings in the literature (Ferrand et al., 1994; Harris et al., 2012; Martin & Barry, 2012; McCloskey, Macaruso, & Rapp, 2006; Sage & Ellis, 2006). Feedback to independent homophone representations facilitating production is potentially the same mechanism that resulted in spoken homophone generalisation of aphasia treatment (e.g., Biedermann, Blanken, & Nickels, 2002; Biedermann & Nickels, 2008a, 2008b) and generalisation to direct neighbours in dysgraphia treatment (Harris et al., 2012; Sage & Ellis, 2006). This assumption is further supported by the finding that homograph but not heterograph priming was observed to be significant, hence, homophone priming was dependent on the degree of orthographic overlap. It appears orthographic overlap has to be very high for priming to occur, either exact overlap (homographs) or with one letter different (direct neighbours). As heterographs are often not direct neighbours (e.g., *flower-flour*), it is possible that there is not sufficient overlap of graphemes between heterographs to facilitate subsequent production.

It is interesting that we found (a suggestion of) priming based on shared orthography, given that the literature on form priming of picture naming has found that this is extremely short lived - it has been found to dissipate after one intervening item (Martin & Barry, 2012; Wheeldon, 2003). In the present study we found orthographic priming with over 100 intervening items suggesting the same mechanism was operating as for repetition priming - that is, priming for orthographically related items via feedback from graphemes but with less feedback and hence less priming than for repetition priming.

One difference between this study and previous studies of written picture naming priming (e.g., Martin & Barry, 2012, Damian et al., 2011) and spoken homophone priming (Wheeldon & Monsell, 1992) is the type of stimuli used. The present study included written picture descriptions beneath our stimuli (to keep similarity between the stimuli used in Chapter 3 and this study), whereas previous picture naming priming studies did not. Comparing the raw reaction times to other written picture naming studies, participants on average took longer to initiate writing the picture name (see Appendix D- 2107.29ms least square mean for unrelated primed items compared to 995ms average primed of unrelated items in Damian et al., 2011). Despite instructing participants to only read the picture description if needed, this increased latency could be caused by the participants only commencing written naming after they had read the picture description. This could diminish the sensitivity of the paradigm to any lexical retrieval effects on target production, and hence, reduce sensitivity to priming.

As our design was a repeated measures, within subjects, design conducted across four blocks, it is entirely possible that, as explained earlier, due to practice effects, subjects became faster at the task with each block. Nevertheless, this was counterbalanced by the fact that half of the 'unprimed' conditions were presented in Session 2 after the same item had been 'primed' in Session 1 (see Table 2). However, this means that there is some possibility that, although Session 2 was conducted at least a week after Session 1, there were some residual priming effects, resulting in the 'unprimed' items being primed by the 'primed' items. This would essentially reduce the priming effect for half of the items. Perhaps the weak homophone and direct neighbour priming effects were a result of this, and a between subjects design where each item was only seen in one condition for each subject may have found more robust priming.

In this study we used latency of initiation of writing as the dependent variable, analogous to speech onset latency in spoken naming. However, writing is known to be a cascading, online process where the process of execution begins when the initial segments have been retrieved (i.e., unlike spoken naming writing is initiated before conflicts are resolved, Delattre, Bonin, & Barry, 2006). Indeed, although writing durations were previously considered to mainly reflect peripheral (execution) processes, recent research has shown lexical effects (i.e., regularity) can influence durations as well as onset of writing (Delattre et al., 2006). Therefore, investigating word durations may provide a more sensitive measure, reflecting not just the onset similarity between prime and target but also similarity of segments at other positions.

In summary, this study shows it is possible to facilitate written picture naming by priming with the same form (repetition and homograph) but also from priming with closely orthographically related forms (direct neighbours). However, only repetition priming produced significantly greater priming than unrelated items. Currently, due to the lack of power, it is unclear, therefore whether the homograph and direct neighbour priming are indeed robust effects or instead artefacts of the methodology used in the present study. Nonetheless, the weak homograph and direct neighbour priming could suggest there is at feedback of activation from prior production in an orthographic system similar in architecture to Middleton et al.'s (2015) Dual Nature account. This suggests priming is more likely to be located in the representations rather than semantic-lexical links. However, future research is needed to confirm these speculative conclusions.

References

Alamargot, D., Chesnet, D., Dansac, C., & Ros, C. (2006). Eye and Pen: A new device for studying reading during writing. *Behavior Research Methods*, *38*(2), 287–299. https://doi.org/10.3758/BF03192780

Alamargot, D., Dansac, C., Chesnet, D., & Fayol, M. (2007). Chapter 2: Parallel Processing Before and After Pauses: A Combined Analysis of Graphomotor and Eye Movements During Procedural Text Production, 11–29. https://doi.org/10.1163/9781849508223_003

Bar, M., & Biederman, I. (1998). Subliminal Visual Priming. *Psychological Science*, *9*(6), 464–468. https://doi.org/10.1111/1467-9280.00086

Barry, C., Hirsh, K. W., Johnston, R. A., & Williams, C. L. (2001). Age of Acquisition, Word Frequency, and the Locus of Repetition Priming of Picture Naming. *Journal of Memory and Language*, *44*(3), 350–375. https://doi.org/10.1006/jmla.2000.2743

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting Linear Mixed-Effects Models using lme4. *arXiv:1406.5823 [Stat]*. Retrieved from http://arxiv.org/abs/1406.5823

Biedermann, B., Blanken, G., & Nickels, L. (2002). The representation of homophones: Evidence from remediation. *Aphasiology*, *16*(10–11), 1115–1136. https://doi.org/10.1080/02687030244000545

Biedermann, B., & Nickels, L. (2008a). Homographic and heterographic homophones in speech production: Does orthography matter? *Cortex*, *44*(6), 683–697. https://doi.org/10.1016/j.cortex.2006.12.001

Biedermann, B., & Nickels, L. (2008b). The representation of homophones: More evidence from the remediation of anomia. *Cortex*, *44*(3), 276–293. https://doi.org/10.1016/j.cortex.2006.07.004

Bonin, P., & Fayol, M. (2002). Frequency effects in the written and spoken production of homophonic picture names. *European Journal of Cognitive Psychology*, *14*(3), 289–313. https://doi.org/10.1080/09541440143000078

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological*, 211–252.

Burke, D. M., Locantore, J. K., Austin, A. A., & Chae, B. (2004). Cherry Pit Primes Brad Pitt Homophone Priming Effects on Young and Older Adults' Production of Proper Names. *Psychological Science*, *15*(3), 164–170. https://doi.org/10.1111/j.0956-7976.2004.01503004.x

Caramazza, A. (1997). How Many Levels of Processing Are There in Lexical Access? *Cognitive Neuropsychology*, *14*(1), 177–208. https://doi.org/10.1080/026432997381664

Caramazza, A., Costa, A., Miozzo, M., & Bi, Y. (2001a). The specific-word frequency effect: Implications for the representation of homophones in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(6), 1430–1450. https://doi.org/10.1037/0278-7393.27.6.1430

Caramazza, A., Costa, A., Miozzo, M., & Bi, Y. (2001b). The specific-word frequency effect: Implications for the representation of homophones in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(6), 1430–1450. https://doi.org/10.1037/0278-7393.27.6.1430

Cave, C. B. (1997). Very Long-Lasting Priming in Picture Naming. *Psychological Science*, *8*(4), 322–325. https://doi.org/10.1111/j.1467-9280.1997.tb00446.x

Damian, M. F., Dorjee, D., & Stadthagen-Gonzalez, H. (2011). Long-term repetition priming in spoken and written word production: Evidence for a contribution of phonology to handwriting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(4), 813–826. https://doi.org/10.1037/a0023260

Davis, C. J. (2005). N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods*, *37*(1), 65–70. https://doi.org/10.3758/BF03206399

Delattre, M., Bonin, P., & Barry, C. (2006). Written spelling to dictation: Sound-to-spelling regularity affects both writing latencies and durations. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *32*(6), 1330–1340. https://doi.org/10.1037/0278-7393.32.6.1330

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, *104*(4), 801–838. https://doi.org/10.1037/0033-295X.104.4.801

Ferrand, L., Grainger, J., & Segui, J. (1994). A study of masked form priming in picture and word naming. *Memory & Cognition*, *22*(4), 431–441. https://doi.org/10.3758/BF03200868

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(4), 680–698. https://doi.org/10.1037/0278-7393.10.4.680

Fry, E. (2004). Phonics: A Large Phoneme - Grapheme Frequency Count Revised. *Journal of Literacy Research*, *36*(1), 85–98. https://doi.org/10.1207/s15548430jlr3601_5

Gordon, J. K. (2002). Phonological neighborhood effects in aphasic speech errors: Spontaneous and structured contexts. *Brain and language*, *82*(2), 113-145.

Harris, L., Olson, A., & Humphreys, G. (2012). Rehabilitation of spelling in a participant with a graphemic buffer impairment: The role of orthographic neighbourhood in remediating the serial position effect. *Neuropsychological Rehabilitation*, *22*(6), 890–919. https://doi.org/10.1080/09602011.2012.709872

Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: experimental and computational studies. *Cognition*, *3*(100), 464–482. https://doi.org/10.1016/j.cognition.2005.02.006

Jacoby, L. L., & Witherspoon, D. (1982). Remembering without awareness. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *36*(2), 300–324. https://doi.org/10.1037/h0080638

Kohnen, S., Nickels, L., Coltheart, M., & Brunsdon, R. (2008). Predicting generalization in the training of irregular-word spelling: treating lexical spelling deficits in a child. *Cognitive Neuropsychology*, *25*(3), 343–375. https://doi.org/10.1080/02643290802003000

Lenth, R. (2016). lsmeans: Least-Squares Means (Version 2.23-5). Retrieved from https://cran.r-project.org/web/packages/lsmeans/index.html

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(01), 1–38.

Martin, D. H., & Barry, C. (2012). Writing nonsense: the interaction between lexical and sublexical knowledge in the priming of nonword spelling. *Psychonomic Bulletin & Review*, *19*(4), 691–698. https://doi.org/10.3758/s13423-012-0261-7

McCloskey, M., Macaruso, P., & Rapp, B. (2006). Grapheme-to-lexeme feedback in the spelling system: Evidence from a dysgraphic patient. *Cognitive Neuropsychology*, *23*(2), 278–307. https://doi.org/10.1080/02643290442000518

Meyer, A. S., Bock, K., (1992). The tip-of-the-tongue phenomenon: Blocking or partial activation? *Memory & Cognition*, *20* (6), 715-726

Middleton, E. L., Chen, Q., & Verkuilen, J. (2015). Friends and foes in the lexicon: Homophone naming in aphasia. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(1), 77–94. https://doi.org/10.1037/a0037778

Mirman, D., Kittredge, A. K., & Dell, G. S. (2010). Effects of near and distant phonological neighbors on picture naming. In *Proceedings of the 32nd annual conference of the Cognitive Science Society* (pp. 1447-1452). Austin, TX: Cognitive Science Society.

Mitchell, D. B., & Brown, A. S. (1988). Persistent repetition priming in picture naming and its dissociation from recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(2), 213–222. https://doi.org/10.1037/0278-7393.14.2.213

Monsell, S., Matthews, G. H., & Miller, D. C. (1992). Repetition of lexicalization across languages: a further test of the locus of priming. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *44*(4), 763–783.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, *76*(2), 165–178. https://doi.org/10.1037/h0027366

Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition*, *114*(2), 227–252. https://doi.org/10.1016/j.cognition.2009.09.007

Perry, C., Ziegler, J. C., & Coltheart, M. (2002). A dissociation between orthographic awareness and spelling production. *Applied Psycholinguistics*, *23*(1), 43–73. https://doi.org/10.1017/S0142716402000036

R Core Team. (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/. Retrieved 17 August 2016, from https://www.google.com.au/search?q=R+Core+Team+(2015).+R%3A+A+language+and+environment+for+statistical+computing.+R+Foundation+for+Statistical+Computing%2C+Vienna%2C+Austria.+URL+https%3A%2F%2Fwww.R-project.org%2F.&ie=utf-8&oe=utf-8&client=firefox-b-ab&gfe_rd=cr&ei=I_C0V6uIGMLr8AeMwrvABw

Sage, K., & Ellis, A. (2006). Using orthographic neighbours to treat a case of graphemic buffer disorder. *Aphasiology*, *20*(9), 851–870. https://doi.org/10.1080/02687030600738945

Vitevitch, M. S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(4), 735.

Wheeldon, L. (2003). Inhibitory form priming of spoken word production. *Language and Cognitive Processes*, *18*(1), 81–109. https://doi.org/10.1080/01690960143000470

Wheeldon, L. R., & Monsell, S. (1992). The Locus of Repetition Priming of Spoken Word Production. *The Quarterly Journal of Experimental Psychology Section A*, *44*(4), 723–761. https://doi.org/10.1080/14640749208401307

White, K. K., Abrams, L., Zoller, S. M., & Gibson, S. M. (2008). Why did I right that? Factors that influence the production of homophone substitution errors. *The Quarterly Journal of Experimental Psychology*, *61*(7), 977–985. https://doi.org/10.1080/17470210801943978

Appendices

Appendix A. Mean frequency (log10) and other psycholinguistic variables by experimental groups

	Repeat (n=19)		Homograph (n=30)		Heterograph (n=16)		Neighbours (n=25)		Unrelated (n=20)	
Variable	A	B	A	B	A	B	A	B	A	B
WWF	1.36	1.36	1.15	1.18	1.44	1.34	0.57	0.66	1.44	1.39
SWF	0.95	0.95	0.85	1.26	0.98	1.08	0.41	0.45	0.93	0.89
WL	1.48	1.48	1.37	1.41	1.54	1.58	0.73	0.78	1.59	1.53
SL	0.99	0.99	0.99	1.21	1.06	1.10	0.49	0.56	1.09	1.05
Syllables	1.05	1.05	1.13	1.13	1.00	1.00	1.17	1.17	1.15	1.20
Phoneme	3.63	3.63	3.00	2.94	3.16	3.24	3.53	3.53	3.35	4.00
Length	4.74	4.74	4.31	4.13	3.96	3.96	4.27	4.27	4.45	4.80
OrthN	4.47	4.47	9.50	6.13	9.76	11.92	8.53	8.53	4.55	4.30
OrthND	290.16	290.16	365.37	29.53	101.80	105.14	114.54	114.54	156.57	50.70
PhonN	11.58	11.58	21.13	22.31	20.12	20.76	17.07	17.07	14.30	11.30
PhonND	258.13	258.13	394.10	394.67	246.96	166.41	106.97	106.97	164.41	242.43
Regular	0.79	0.89	0.50	0.50	0.88	0.92	0.90	0.90	0.60	0.65

*=p<.05 in independent sample t-tests between set A and B for each experimental set
WWF= written word form frequency
SWF= spoken word form frequency
WL= written lemma frequency
SL= spoken lemma frequency
Orth N= orthographic neighbourhood
OrthND= orthographic neighbourhood density
PhonN= phonological neighbourhood
PhonND= phonological neighbourhood density

Appendix B. The Items In Each Set Used Both As Primes And Targets

Repeat	Homograph		Heterograph		Direct Neighbour		Unrelated	
	A	B	A	B	A	B	A	B
Ant	Bank(River)	Bank(Money)	Cue	Queue	Belt	Bell	Arch	Cliff
Bench	Bark(Tree)	Bark(Dog)	Dough	Doe	Bomb	Comb	Arm	Ghost
Branch	Bat(Animal)	Bat(Hit)	Flower	Flour	Bone	Cone	Axe	Thorn
Bread	Boot(Shoe)	Boot	Hare	Hair	Brain	Drain	Bed	Thumb
Chalk	Bow(Arrow)	Bow(Ribbon)	Leak	Leek	Cake	Cave	Bus	Tent
Cheese	Bulb(light)	Bulb(Flower)	Lock	Loch	Clown	Crown	Chief	Desk
Church	Calf(Leg)	Calf(Cow)	Mousse	Moose	Coal	Coat	Dew	Knee
Crab	Chest(Body)	Chest(gold)	Muscle	Mussel*	Cork	Fork	Flag	Neck
Face	China(Plate)	China(World)	Night	Knight	Cot	Cow	Knife	Brush
Fig*	Court(Jury)	Court(Tennis)	Peace	Piece	Ear	Car	Lake	List
Milk	Dummy(Crash)	Dummy(Baby)	Pear	Pair	Hall	Hill	Lemon	Coffee
Moon	Fan(Football)	Fan(Cool)	Pie	Pi	Ham	Ram	Pearl	Fish
Rake	Glass(Window)	Glass(Water)	Root	Route	Hand	Sand	Phone	Sleeve
Skirt	Horn(Brass)	Horn(Nose)	Sail	Sale	Jug	Rug	Sheep	Tulip
Skull	Letter(ABC)*	Letter(Pen)	Son	Sun	King	Wing	Star	Dress
Snail	Mole(Animal)	Mole(Skin)	Tee	Tea	Lamb	Lamp	Sugar	Ambulance
Taxi	Mouse(Comp)	Mouse(rat)			Leg	Log	Thigh	Gold
Throne	Nail(Finger)	Nail(Hammer)			Lime	Line	Throat	Train
Tongue	Nut(Bolt)	Nut(Eat)			Path	Bath	Trousers	Swan
	Palm(Hand)	Palm(Tree)			Pink	Sink	Vase	Mop
	Pen(Write)	Pen(Pigs)			Plate	Slate		
	Pipe(Copper)	Pipe(Smoke)			Shade	Spade		
	Plug(Bath)	Plug(Electric)			Shed	Seed		
	Second(2nd)	Second(Time)			Shoe	Shop		
	Suit(Card)	Suit(Tie)			Soap	Soup		
	Tap(Dance)	Tap(Water)						
	Temple(Head)	Temple(Church)						
	Toast(Jam)	Toast(Drink)						
	Trunk(Nose)	Trunk(Tree)						
	Wave(Sea)	Wave(Hand)						

* Excluded From Analysis Due To <40% Spelling Agreement Along With Partner

Appendix C. The various models tested with the *p* values and chi squared (Chisq) values from ANOVAs comparing each model with the previously significant model using the transformed onset reaction times for written picture naming (significant models highlighted in bold)

	Model	Chisq	p
1	(1|subject)+(1|items)		
2	(group)+(1|subject)+(1|item)	34.09	<.001
3	(group+prime)+(1|subject)+(1|item)	31.02	<.001
4	(group*prime)+(1|subject)+(1|item) †	6.90	.28
5	(group*prime)+SWF+(1|subject)+(1|Item)	0.04	.83
6	(group*prime)+ WWF+(1|subject)+(1|Item)	0 .00	1.00
7	group*prime)+WL+(1|subject)+(1|Item	0.67	.41
8	(group*prime) +SL+(1|subject)+(1|Item)	0.06	.81
9	(group*prime)+ length+(1|subject)+(1|Item)	2.55	.11
10	(group*prime)+syllables+(1|subject)+(1|Item)	6.35	.01
11	(group*prime)+syllables+phoneme+(1|subject)+(1|Item)	0.57	.45
12	(group*prime)+syllables+orthN+(1|subject)+(1|Item),	3.65	.06
13	(group*prime) +syllables+orthND+(1|subject)+(1|Item)	0.25	.62
14	(group*prime)+ syllables+orderC+(1|subject)+(1|Item),	0.43	.51
15	(group*prime)+syllables+presentationOrder+(1|subject)+(1|Item),	204.53	<.001
16	group*prime)+syllables+presentationOrder+NA+(1|subject)+(1|Item)	32.51	<.001

group= repeat, homograph, heterography, repeat, unrelated
prime- unprimed, primed
SWF= spoken word form frequency (centred)
WL= written lemma frequency (centred)
SL= Spoken lemma frequency (centred)
WWF= written word form frequency (centred)
length= number of letters(centred)
syllables= number of syllables(centred)
Phoneme= number of phonemes(centred)
OrthN= number of orthographic neighbours(centred)
OrthND= average frequency of orthographic neighbours(centred)
OrderC= the presentation order within the block (centred)
presentationOrder= if the prime was presented before or after the target
NA= name agreement, percentage of correct name produced(cantered)

† Although adding the interaction of group by prime did not significantly increase the model fit, we have strong a priori reasons to assume this would be the case, as the true effects are masked within the group fixed factor and hidden by practice effects over the repeated measures sessions. To investigate this with contrast coding this interaction needs to be included in the linear mixed effects

Appendix D. The raw least squared means RTs and standard error in brackets for the groups by prime

	Unprimed		Primed	
Repeat	1962.47	85.58	1828.94	85.89
Homograph	2159.84	79.06	2051.47	79.80
Heterograph	2118.29	90.86	2087.89	92.54
Direct	1978.31	80.27	1907.99	80.87
Unrelated	2123.67	82.13	2107.29	82.86

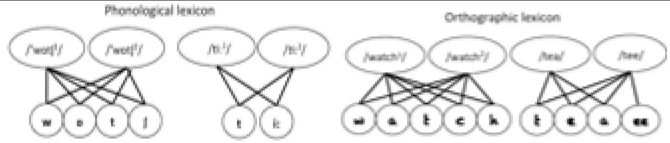
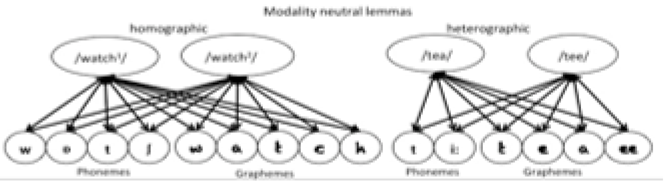
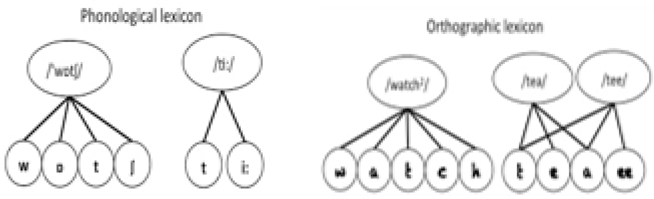
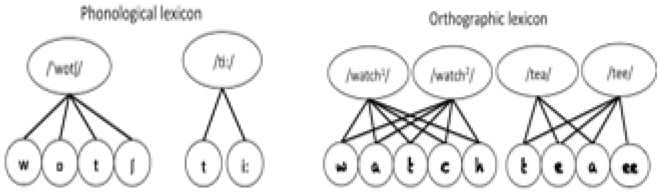
Chapter 5:

Summary and Conclusions: Thyme for a knew perspective?

SUMMARY AND CONCLUSIONS

The aim of this thesis was to investigate homophone representation in the orthographic and phonological output lexicons. My particular interest comprised the processing steps taken from the semantic system to the activation and selection of phonemes or letters in order to produce a word. Currently, there are three major theories in the spoken language production literature. The first theory proposes that homophones have separate lemma (lexical syntactic) representations that feed-forward to a shared word form level (as proposed in the Two-Stage model, The second proposes that homophones have independent word form representations that are accessed in a cascading fashion alongside optional but not obligatory lexical-syntactic activation (as proposed in the Independent Network model,. The third model postulates that homophones have independent lemmas which compete for selection, influenced by interactive feedback links from phonemes (as proposed in the Dual Nature account, Middleton, Chen, & Verkuilen, 2015). These three models all make different predictions as to how homophones are processed. A summary of the models and the different predictions they make was presented in Chapter 1 of the thesis and is reproduced in Table 1 below.

Table 1. The possible representations of homographs and heterographs in the phonological and orthographic lexicon and their predicted effects.

	Representation	Predicted Advantage?			
		Spoken		Written	
		Homo	Het	Homo	Het
1	 Independent in both lexicons and feedforward activation e.g. Independent Network	✗	✗	✗	✗
2	 Independent in both lexicons with feedback and competition e.g. Dual nature account	✓/✗	✓/✗	✓/✗	?
3	 Shared in phonological lexicon, only homographs shared in orthographic and feedforward activation e.g., The most likely version of the Two-Stage model adapted to include the orthographic lexicon	✓	✓	✓	✗
4	 Shared in phonological lexicon only and feedforward activation e.g. an alternative adaptation of the Two-Stage model to include the orthographic lexicon	✓	✓	✗	✗

Homo= homograph
Het= heterograph
✓ Advantage e.g. frequency inheritance, generalisation of homophone treatment, homophone priming
✗ No advantage e.g. homophones produced as individual frequency would suggest, no generalisation and priming or only to the extent predicted by segment overlap/feedback,
✓/✗ Both advantage and no advantage depending on the stage of retrieval

The Two-Stage model predicts a homophone advantage (e.g., faster reaction times than the individual word form frequency predicts, known as the frequency inheritance effect established by Jescheniak & Levelt, 1994; or generalisation of treatment effects to untreated homophone partners, Biedermann and colleagues, 2002; 2008a, 2008,b). The Independent Network model predicts no homophone advantage (e.g., homophone latencies according to individual homophone frequency, Caramazza et al., 2001), whereas the Dual Nature account (Middleton et al., 2015, and based on Dell and colleagues Interactive Activation model, e.g. Dell, Schwartz, Martin, Saffran, & Gagnon, 1997) has more complex predictions. For selection of lemmas (Stage 1 retrieval) there may be either a disadvantage or no advantage depending on the task, while for selection of phonemes (Stage 2 retrieval), and particularly in tasks with less semantic involvement, there is likely to be a homophone advantage (Middleton et al., 2015).

A particular aim of this thesis was to investigate whether differences in homophone production across the phonological and orthographic lexicons interacted with spelling of homophones (homographic homophones, e.g., *nut*- food and *nut*-bolt vs. heterographic homophones e.g., *route*-*root*). I was also interested in which of the spoken theories of homophone representation mentioned above would best describe written homophone production.

In this concluding chapter, I will summarise the aims and results of the three experimental chapters before drawing together my findings across all three studies within the three theoretical frameworks mentioned above.

Summary of Experimental Chapters

Chapter 2: Two bee or not to be: The effects of orthography and bilingualism on spoken homophone production.

Chapter 2 investigated spoken homophone production in unimpaired monolingual and bilingual picture naming and (bilingual) translation in two experiments. The aim of this study was to investigate if previous contrasting findings were caused by differences in participants (monolingual or bilingual), spelling (heterographic or homographic) or task (picture naming or translation).

The picture naming study (Experiment 1) found no advantage for homophone production compared to frequency matched controls. Instead, it was found that homophones were produced at similar speed to individual frequency matched controls and significantly slower than controls matched to the summed frequency of both homophone meanings. The same effects were found across participants and across the homophone subsets (heterographs and homographs) in this picture naming task. This experiment therefore allows us to conclude that previously conflicting evidence cannot be attributed to differences in participants' language background (whether they were monolingual or bilingual). Similarly, naming pictures of homophones resulted in the same effect (no advantage) for heterographs and homographs, suggesting that homophone spelling is probably not the cause for inconsistent results in previous studies.

The translation task (Experiment 2) also resulted in no advantage for homographs compared to non-homophones, but the pattern for heterographs was less clear. However, the fact we did not match psycholinguistic variables cross-linguistically (only the English targets were matched but not the Welsh to-be-translated stimuli) means that this difference is hard to interpret. Hence, while this study gives some suggestion that task (i.e., picture naming or translation) may influence the patterns found, the results from this study alone are not decisive and therefore warrant further investigation.

Chapter 3. Too harts won sole: Dysgraphia treatment outcomes as evidence for homophone representation**.**

The second experimental chapter investigated generalisation of dysgraphia treatment to homophones and direct neighbours with the aim of understanding i) if orthographic homophone treatment shows the same pattern of generalisation to the untreated partner previously found for the spoken modality (see Biedermann and colleagues, 2002; 2008a, 2008,b), and ii) if generalisation was found for homophone spelling, was it caused by feedback or a shared word form representation?

Although the treatment resulted in improvement for treated items, no generalisation to homophones was found. We further found no generalisation to direct neighbours. While a lack of generalisation to direct neighbours has also been observed in other dysgraphia treatment studies (e.g., Krajenbrink, Nickels, & Kohnen, 2016), our findings contradict previous studies that do report significant generalisation to untreated neighbours (Harris et al., 2012; Sage & Ellis, 2006).

Unfortunately, this study is unable to determine if the absence of generalisation was caused by the nature of our participant's impairment resulting in no effect of feedback from graphemes to orthographic word form representations, or if feedback is absent in orthographic production altogether. The latter argument seems unlikely since there is a growing body of research that finds evidence for feedback from graphemes to orthographic word form representations (e.g.,Harris et al., 2012; McCloskey, Macaruso, & Rapp, 2006; Sage & Ellis, 2006). Nevertheless, the results of this study clearly support independent homophone representation(s) in the orthographic lexicon, but cannot discriminate between accounts that do or do not implement feedback: further investigation is needed to tackle the origin of the observed lack of feedback.

Chapter 4. Homophone priming of written picture naming: Witch root for prior righting?

In analogy to the treatment study (Chapter 3), the final experimental study (Chapter 4) investigated homophone priming. It aimed to investigate whether homophones share a representation at the orthographic word form level, and the level at which priming has its effect. The role of feedback from graphemes to the lexicon was also investigated by priming orthographic direct neighbours. This study sought to better understand the mechanism that might have caused the lack of feedback found in Chapter 3. Unfortunately, the bespoke software required to extract the data caused unforeseen technical problems that could not be solved in the remaining time frame available for this thesis. Hence, this study should be considered pilot data and consequently it is difficult to draw firm conclusions from this experiment. Nonetheless, we found significant priming from identity primes – naming a picture was faster when that item (but not that same picture – e.g., two different types of 'dogs' are depicted) had been named previously than when it had not previously been named. This orthographic repetition priming supports previous findings (Damian et al., 2011; Martin & Barry, 2012). In addition, like previous spoken priming (Wheeldon & Monsell, 1992), we found some suggestion of significant priming for homographs, but not for heterographs. Similar effects were found for direct neighbours. Together these results suggest that priming has its effects at the level of word form representations (which are lemma to segment (phoneme/grapheme) links in the Dual Nature account) and that homograph priming results from feedback from shared graphemes to the independent orthographic word forms. Of course, further research (and/or more data) is required to confirm this hypothesis.

Theoretical implications

In sum, Chapter 2 and Chapter 3 suggest no advantage for homophone production in either spoken or written modalities. This is broadly supported by Chapter 4, but with some suggestion that there may be an advantage for homophones dependent on orthographic overlap. These findings will be discussed in terms of the three theories of homophone representation.

Shared representations in the lexicon

The results from the three studies in this thesis are not compatible with the shared homophone word form assumption of the Two-Stage model (Levelt et al., 1999). If homophones were shared at the phonological and/or orthographic word form level, we would expect to find a homophone advantage across all three studies (see Table 1 Rows 3 and/or 4). Instead, it was found that homophones were produced similarly to non-homophonic controls (no advantage) in Chapters 2 and 3, and had no strong evidence of an advantage in the priming study (Chapter 4). The combination of the results across all the chapters, plus many previous findings (e.g., Bonin & Fayol, 2002; Caramazza et al., 2001), together argue against homophones having shared representations at the word form level.

Independent representations in the lexicon

The pattern of results across the three experimental chapters most strongly support homophones having independent representations at the word form level, as specified in the Independent Network model (Caramazza et al., 2001), or in the Dual Nature account (Middleton et al., 2015- Table 1 Row 2). Although both these models include independent representations for homophones, the Independent Network model considers activation to be feedforward only, and the independent representations to be modality-specific (i.e., each item has one phonological representation and one orthographic representation). The Dual Nature account, on the other hand, assumes interactive activation, whereby activation from graphemes or phonemes feeds back to the representations, which are modality-neutral.

The differences between these two models result in different predictions. The Independent Network model predicts no difference between production of homophones and non-homophones (see Table 1 Row 1): this is what we find in Chapter 3 and Chapter 2, Experiment 1 where there was no generalisation of homophone treatment, and homophones were named as quickly as frequency matched controls. However, the Independent Network model has difficulty explaining the partial advantage found for some homophones in translation (Chapter 2) and priming (Chapter 4). It also has difficulty explaining any difference between homographs and heterographs (as was found in the translation task, Chapter 2 Experiment 2). This is because homophone representation is identical regardless of spelling – they are represented independently while being modality-specific. Moreover, without feedback there is no way that orthography can influence phonology. The Independent Network model predicts that previous priming by a homograph or an orthographically related item would have no effect on subsequent production, however, there was some suggestion of such effects in Chapter 4.

On the other hand, the Dual Nature account predicts that the extent of a homophone advantage can vary dependent on the influence of each stage of retrieval (see Table 1 Row 2). In Chapter 2, we found no advantage in homophone picture naming, which is predicted by the Dual Nature account as it is a Stage 1 driven task. In comparison, tasks with less semantic emphasis (Stage 2 driven tasks) may show an advantage. While not significant, there was some indication of an advantage in translating heterographic homophones. As the Dual Nature account contains modality-neutral (lemma) representations, this explains how, with feedback, orthography can influence spoken word production resulting in differing effects for heterographs and homographs. I suggested that while all homophones would be activated by reverberating activation from shared phonemes perhaps the balance between competition and increased activation tipped towards an advantage for heterographs due to their reduced overlap. Clearly, this is speculative and simulation is required to determine the relative effects of different factors in predicting reaction times.

This account also, most probably, predicts generalisation in treatment. While this was not found in Chapter 3, as noted above, this could have been due to the nature of the patient's impairment (potentially an additional buffer impairment reducing feedback) and so cannot help us adjudicate on the lexical architecture.

The suggestion of priming effects related to shared orthography found in Chapter 4 can be explained by the interactive nature of the Dual Nature account. Prior production of an item will result in strengthening of the links between that item's lemma and its graphemes. Consequently, when a subsequent item that shares graphemes is activated, reverberation of activation will result in greater activation of those graphemes shared between prime and target. Therefore, the greater the overlap of orthography is, the greater the benefit to production will be.

To summarise, while the definitive results of this thesis are consistent with the Independent Network model, across the experiments there are suggestions of patterns that support feedback influencing homophone production. Increasingly, the broader evidence base in the literature also supports feedback from graphemes and/or phonemes to word form representations. Therefore the inclusion of feedback from graphemes to the lexical level appears warranted. Consequently, it seems that the Dual Nature account (Table 1 Row 2), provides the most plausible explanation of homophone representation, according to the results found in this thesis and the previous findings in the literature. The Dual Nature model has the benefit of having the potential to explain both an advantage for homophones (e.g., Chapter 4, Biedermann et al., 2002; Biedermann & Nickels, 2008b, 2008a; Biran, Gvion, Sharabi, & Gil, 2013; Jescheniak & Levelt, 1994; Middleton et al., 2015) and no advantage (Chapter 2-Experiment 1, e.g., Bonin & Fayol, 2002; Caramazza et al., 2001; Jacobs, Singer, & Miozzo, 2004), hence seems the most appealing theory given the reality of an 'inconsistent' evidence base.

Nevertheless, further research into the Dual Nature Account of homophony is required. The focus of this account to date has been to simulate performance of people with aphasia and different levels of impairment (Middleton et al., 2015). What would be advantageous now is to run simulations of homophone production under different conditions (e.g., priming, different tasks). Such simulation is vital in order to confirm the hypotheses suggested by Middleton et al. (2015) that were further explored in this thesis, and hence provide further testable hypotheses.

Conclusions

The aim of this thesis was to resolve and disentangle the conflict regarding investigations of homophone representation. The strongest pattern in this thesis was no apparent advantage for homophones, suggesting that homophones have independent representations in both lexicons. Nevertheless, there were suggestions of a potential advantage under some conditions, which taken together with other research, advocates for the Dual Nature account (Middleton et al., 2015) as currently the most promising model of homophone representation due to its ability to explain varied degrees of homophone (dis)advantage in production. Nevertheless, further research is needed to fully explore this novel account.

In sum, through a convergence of research methods, this thesis has systematically explored the nature of homophone representation and processing, giving further insights into the complexities of this unique lexical class and providing pointers for future research efforts.

References

Biedermann, B., Blanken, G., & Nickels, L. (2002). The representation of homophones: Evidence from remediation. *Aphasiology*, *16*(10–11), 1115–1136. https://doi.org/10.1080/02687030244000545

Biedermann, B., & Nickels, L. (2008a). Homographic and heterographic homophones in speech production: Does orthography matter? *Cortex*, *44*(6), 683–697. https://doi.org/10.1016/j.cortex.2006.12.001

Biedermann, B., & Nickels, L. (2008b). The representation of homophones: More evidence from the remediation of anomia. *Cortex*, *44*(3), 276–293. https://doi.org/10.1016/j.cortex.2006.07.004

Biran, M., Gvion, A., Sharabi, L., & Gil, M. (2013). The Representation of Homophones in the Phonological Lexicon: Additional Evidence from Hebrew. *Procedia - Social and Behavioral Sciences*, *94*, 86–87. https://doi.org/10.1016/j.sbspro.2013.09.041

Bonin, P., & Fayol, M. (2002). Frequency effects in the written and spoken production of homophonic picture names. *European Journal of Cognitive Psychology*, *14*(3), 289–313. https://doi.org/10.1080/09541440143000078

Bonin, P., Méot, A., Lagarrigue, A., & Roux, S. (2015). Written object naming, spelling to dictation, and immediate copying: Different tasks, different pathways? *The Quarterly Journal of Experimental Psychology*, *68*(7), 1268–1294. https://doi.org/10.1080/17470218.2014.978877

Caramazza, A., Costa, A., Miozzo, M., & Bi, Y. (2001). The specific-word frequency effect: Implications for the representation of homophones in speech production. *Journal* *of Experimental Psychology: Learning, Memory, and Cognition*, *27*(6), 1430–1450. https://doi.org/10.1037/0278-7393.27.6.1430

Damian, M. F., Dorjee, D., & Stadthagen-Gonzalez, H. (2011). Long-term repetition priming in spoken and written word production: Evidence for a contribution of phonology to handwriting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(4), 813–826. https://doi.org/10.1037/a0023260

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, *104*(4), 801–838. https://doi.org/10.1037/0033-295X.104.4.801

Harris, L., Olson, A., & Humphreys, G. (2012). Rehabilitation of spelling in a participant with a graphemic buffer impairment: The role of orthographic neighbourhood in remediating the serial position effect. *Neuropsychological Rehabilitation*, *22*(6), 890–919. https://doi.org/10.1080/09602011.2012.709872

Hughes, E. K., Roberts, R. J., Roberts, D. J., Kendrick, L. T., Payne, J. S., Owen-Booth, B., … Tainturier, M.-J. (2014). Is translation semantically mediated? Evidence from Welsh-English bilingual aphasia. *Frontiers in Psychology: Academy of Aphasia-52nd Annual Meeting*. https://doi.org/10.3389/conf.fpsyg.2014.64.00082

Jacobs, M., Singer, N., & Miozzo, M. (2004). The representation of homophones: Evidence from anomia. *Cognitive Neuropsychology*, *21*(8), 840–866. https://doi.org/10.1080/02643290342000573

Jescheniak, J. D., & Levelt, M., J. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 824–843. https://doi.org/10.1037/0278-7393.20.4.824

Krajenbrink, T., Nickels, L., & Kohnen, S. (2016). No effect of orthographic neighbourhood in treatment with two cases of acquired dysgraphia. *Aphasiology*, *0*(0), 1–34. https://doi.org/10.1080/02687038.2016.1218435

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(01), 1–38.

Martin, D. H., & Barry, C. (2012). Writing nonsense: the interaction between lexical and sublexical knowledge in the priming of nonword spelling. *Psychonomic Bulletin & Review*, *19*(4), 691–698. https://doi.org/10.3758/s13423-012-0261-7

McCloskey, M., Macaruso, P., & Rapp, B. (2006). Grapheme-to-lexeme feedback in the spelling system: Evidence from a dysgraphic patient. *Cognitive Neuropsychology*, *23*(2), 278–307. https://doi.org/10.1080/02643290442000518

Middleton, E. L., Chen, Q., & Verkuilen, J. (2015). Friends and foes in the lexicon: Homophone naming in aphasia. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(1), 77–94. https://doi.org/10.1037/a0037778

Sage, K., & Ellis, A. (2006). Using orthographic neighbours to treat a case of graphemic buffer disorder. *Aphasiology*, *20*(9), 851–870. https://doi.org/10.1080/02687030600738945

Wheeldon, L. R., & Monsell, S. (1992). The Locus of Repetition Priming of Spoken Word Production. *The Quarterly Journal of Experimental Psychology Section A*, *44*(4), 723–761. https://doi.org/10.1080/14640749208401307

Appendix
Ethical Approval

Macquarie University

MACQUARIE UNIVERSITY

Lyndsey Nickels <lyndsey.nickels@mq.edu.au>

Approved- Ethics application- Nickels (Ref No: 5201200905)

Ethics Secretariat <ethics.secretariat@mq.edu.au> 12 December 2012 11:17
To: Prof Lyndsey Nickels <lyndsey.nickels@mq.edu.au>

Dear Prof Nickels

Re: "Understanding language processing, its breakdown and treatment"
(Ethics Ref: 5201200905)

Thank you for your recent correspondence. Your response has addressed the issues raised by the Human Research Ethics Committee and you may now commence your research.

This research meets the requirements of the National Statement on Ethical Conduct in Human Research (2007). The National Statement is available at the following web site:

http://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/e72.pdf.

The following personnel are authorised to conduct this research:

Dr Anna Elisabeth Beyersmann
Dr Antje Lorenz
Dr Britta Biedermann
Dr Karen Croot
Dr Kati Renvall
Dr Samantha Siyambalapitiya
Dr Saskia Kohnen
Miss Anastasiia Romanova
Miss Catherine Mason
Miss Trudy Geertruida Krajenbrink
Ms Belinda McDonald
Ms Emily Church
Ms Fransizka Bachmann
Ms Jennifer Cole-Virtue
Ms Nora Fieder
Polly Barr
Prof Lyndsey Nickels
Prof Niels Schiller
Professor David Howard
Rimke Groenewold
Shiree Heath
Solene Hameau
Tina Marusch

NB. STUDENTS: IT IS YOUR RESPONSIBILITY TO KEEP A COPY OF THIS APPROVAL EMAIL TO SUBMIT WITH YOUR THESIS.

Please note the following standard requirements of approval:

1. The approval of this project is conditional upon your continuing compliance with the National Statement on Ethical Conduct in Human Research (2007).

Progress Report 1 Due: 12 December 2013
Progress Report 2 Due: 12 December 2014
Progress Report 3 Due: 12 December 2015
Progress Report 4 Due: 12 December 2016
Final Report Due: 12 December 2017

NB. If you complete the work earlier than you had planned you must submit a Final Report as soon as the work is completed. If the project has been discontinued or not commenced for any reason, you are also required to submit a Final Report for the project.

Progress reports and Final Reports are available at the following website:

http://www.research.mq.edu.au/for/researchers/how_to_obtain_ethics_approval/human_research_ethics/forms

3. If the project has run for more than five (5) years you cannot renew approval for the project. You will need to complete and submit a Final Report and submit a new application for the project. (The five year limit on renewal of approvals allows the Committee to fully re-review research in an environment where legislation, guidelines and requirements are continually changing, for example, new child protection and privacy laws).

4. All amendments to the project must be reviewed and approved by the Committee before implementation. Please complete and submit a Request for Amendment Form available at the following website:

http://www.research.mq.edu.au/for/researchers/how_to_obtain_ethics_approval/human_research_ethics/forms

5. Please notify the Committee immediately in the event of any adverse effects on participants or of any unforeseen events that affect the continued ethical acceptability of the project.

6. At all times you are responsible for the ethical conduct of your research in accordance with the guidelines established by the University. This information is available at the following websites:

http://www.mq.edu.au/policy/

http://www.research.mq.edu.au/for/researchers/how_to_obtain_ethics_approval/human_research_ethics/policy

If you will be applying for or have applied for internal or external funding for the above project it is your responsibility to provide the Macquarie University's Research Grants Management Assistant with a copy of this email as soon as possible. Internal and External funding agencies will not be informed that you have final approval for your project and funds will not be released until the Research Grants Management Assistant has received a copy of this email.

Please retain a copy of this email as this is your official notification of final ethics approval.

Yours sincerely
Dr Karolyn White
Director of Research Ethics
Chair, Human Research Ethics Committee

Bangor University

GIG CYMRU NHS WALES | Bwrdd Iechyd Prifysgol Betsi Cadwaladr University Health Board

Panel Arolygu Mewnol Y&D – Y Gorllewin
R&D Internal Review Panel – West Division

Ysbyty Gwynedd
Clinical Academic Office
North Wales Clinical School
Bangor, Gwynedd
LL57 2PW

Telephone/Facsimile: 01248 - 384.877
Email: Rossela.Roberts@nww-tr.wales.nhs.uk

PRIVATE & CONFIDENTIAL
Dr Marie-Josephe Tainturier
School of Psychology
Brigantia Building
Bangor Univeristy
Bangor
LL57 2AS

21 April 2011

Dear Dr Tainturier,

Re: Project Review:

Tainturier 10/WNo01/67 Patterns of crosslinguistic treatment generalisation in acquired language disorders: A window into the organisation of the bilingual language system

The above research project was reviewed at the meeting of the Internal Review Panel held on 6 January 2011. Thank you for responding to the Committee's request for further information.

The Chairman considered the response on behalf of the Committee and is satisfied with the scientific validity of the project, the risk assessment, the review of the NHS cost and resource implications and all other research management issues pertaining to the revised application.

I have pleasure in confirming that the Internal Review Panel is pleased to grant approval to proceed at Betsi Cadwaladr University Health Board (West sites).

The study should not commence until the Ethics Committee reviewing the research has confirmed final ethical approval - favourable opinion.

All research conducted at the Betsi Cadwaladr University Health Board sites must comply with the Research Governance Framework for Health and Social Care in Wales (August 2009).
An electronic link to this document is provided on the Trust's R&D WebPages.
Alternatively, you may obtain a paper copy of this document via the R&D Office.

Attached you will find a set of approval conditions outlining your responsibilities during the course of this research. Failure to comply with the approval conditions will result in the withdrawal of the approval to conduct this research in the Betsi Cadwaladr University Health Board.

If you would like further information on any other points covered by this letter please do not hesitate to contact me. On behalf of the Committee, may I take this opportunity to wish you every success with your research.

Chairman/Cadeirydd – Dr Richard Tranter

GIG CYMRU NHS WALES
Bwrdd Iechyd Prifysgol
Betsi Cadwaladr
University Health Board

Panel Arolygu Mewnol Y&D – Y Gorllewin
R&D Internal Review Panel – West Division

Ysbyty Gwynedd
Clinical Academic Office
North Wales Clinical School
Bangor, Gwynedd
LL57 2PW

Telephone/Facsimile: 01248 - 384.877
Email: Rossela.Roberts@nww-tr.wales.nhs.uk

List of documents reviewed and approved:

Documents reviewed:	Version	Date
NHS R&D Form	-	17/11/2010
Protocol	1.1	14/02/2011
Patient Information Sheet (MRI)	1.0	06/12/2010
Consent Form (MRI)	1.0	06/12/2010
Safety Checklist (MRI)	1.0	06/12/2010
Debriefing Form (MRI)	1.0	06/12/2010
Debriefing Form (Phase 2)	1.0	06/12/2010
Information Sheet (Phase 1)	1.1	18/04/2011
Information Sheet (Phase 2)	1.1	18/04/2011
Consent Form (Phase 1)	1.1	18/04/2011
Consent Form (Phase 2)	1.1	18/04/2011
GP Letter	1.0	23/02/2011
Project Partner Letter of Support		25/06/2009

Yours sincerely

pp

Dr Richard Tranter
Consultant Psychiatrist, Associate Director of R&D
Chairman R&D Internal Review Panel

Chairman/Cadeirydd – Dr Richard Tranter

NHS REC Form | Reference: 10/WNo01/61 | IRAS Version 3.1

members of the Chief Investigator's team, including non-doctoral student researchers.

	Title Forename/Initials Surname
	Dr Daniel Roberts
Post	Research Officer
Qualifications	PhD
Employer	Bangor University
Work Address	School of Psychology Brigantia Building Bangor
Post Code	LL57 2AS
Telephone	07793707989
Fax	01248382599
Mobile	07793707989
Work Email	d.roberts@bangor.ac.uk

	Title Forename/Initials Surname
	Dr Charles Leek
Post	Reader
Qualifications	PhD
Employer	Bangor University
Work Address	School of Psychology Brigantia Building Bangor
Post Code	LL57 2AS
Telephone	01248382948
Fax	01248382599
Mobile	
Work Email	e.c.leek@bangor.ac.uk

	Title Forename/Initials Surname
	Prof Virginia Gathercole
Post	Co-Director, ESRC Centre for Research on Bilingualism in Theory and Practice
Qualifications	PhD
Employer	Bangor University
Work Address	School of Psychology Brigantia Building Bangor
Post Code	LL57 2AS
Telephone	01248382624
Fax	01248382599
Mobile	
Work Email	v.c.gathercole@bangor.ac.uk

	Title Forename/Initials Surname
	Miss Emma Hughes
Post	PhD student
Qualifications	MRes

Date: 08/11/2010 18 66308/171555/1/518

Work Email	psp869@bangor.ac.uk

	Title Forename/Initials Surname Miss Polly Barr
Post	Research Project Support Officer
Qualifications	BSc
Employer	Bangor University
Work Address	School of Psychology Brigantia Building Bangor
Post Code	LL57 2AS
Telephone	01248383010
Fax	
Mobile	
Work Email	p.barr@bangor.ac.uk

	Title Forename/Initials Surname Prof Robert Rafal
Post	Consultant Neurologist
Qualifications	PhD
Employer	Bangor University
Work Address	School of Psychology Brigantia Building Bangor
Post Code	LL57 2AS
Telephone	01248383603
Fax	01248382599
Mobile	
Work Email	r.rafal@bangor.ac.uk

	Title Forename/Initials Surname Dr Giovanni d'Avossa
Post	Neurologist
Qualifications	PhD
Employer	Bangor University
Work Address	School of Psychology Brigantia Building Bangor
Post Code	LL57 2AS
Telephone	01248388801
Fax	01248382599
Mobile	
Work Email	g.davossa@bangor.ac.uk

	Title Forename/Initials Surname Dr Martyn Bracewell
Post	Neurologist
Qualifications	PhD
Employer	Bangor University