Running Head: HOMOPHONE SUBSTITUTION ERRORS

Why Did I Right That? Factors that Influence the Production of Homophone Substitution Errors

Katherine K. White
College of Charleston

Lise Abrams
University of Florida

Sarah M. Zoller
College of Charleston

Samantha M. Gibson
College of Charleston

Please address correspondence to:

Dr. Katherine White
Department of Psychology
College of Charleston
57 Coming Street
Charleston, SC 29424

Phone: (843) 953-5517
Fax: (843) 953-7151
Email: whitek@cofc.edu

WORD COUNT: 3,792
Abstract

Despite considerable research on language production errors involving speech, little research exists in the complementary domain of writing. Two experiments investigated the production of homophone substitution errors, which occur when a contextually appropriate word (e.g., *beech*) is replaced with its homophone, e.g., *beach* tree. Participants wrote down auditorily-presented sentences containing dominant or subordinate homophones. Homophones were preceded by a lexical prime that overlapped in phonology and orthography (e.g., *teacher*) or only orthography (e.g., *headmaster*) with the target homophone. Results showed more substitution errors when the context elicited a subordinate homophone relative to a dominant homophone. Furthermore, both types of primes equivalently increased production of homophone errors relative to control word (e.g., *lawyer*), suggesting that only orthographic overlap between the prime and target was necessary to influence errors. These results are explained within dual route models of spelling, which postulate an interaction between lexical and sublexical routes when spelling.
Why Did I Right That? Factors that Influence the Production of Homophone Substitution Errors

Speech errors are widely used as tools to understand the processes underlying spoken language production and to investigate the roles of phonology, morphology, semantics, and syntax in production (e.g., Dell, 1986; Levelt, Roelofs, & Meyer, 1999; Vitevitch, 2002). Although considerable research has been conducted on speech errors, or “slips of the tongue”, little research has investigated comparable errors in written language production, or “slips of the pen” (e.g., Hotopf, 1983). The present study extended research on language production errors to slips of the pen that occur when homophones (e.g., beach, beech) substitute for one another in written contexts. Unlike previous research investigating the role of morphological errors in the homophone spelling process (e.g., Largy, Fayol, & Lemaire, 1996), we used a written error induction paradigm to examine orthographically distinct homophone substitutions. Specifically, we investigated the influence of phonological and orthographic primes, as well as homophone dominance, on substitution errors.

Dual route models of spelling offer a framework to make predictions about the sources of substitution errors because of their depiction of the processes underlying spelling retrieval from memory as well as assembled spellings from phoneme-to-grapheme (P-G) mappings (e.g., Barry, 1994; Houghton & Zorzi, 2003; Rapp, Epstein, & Tainturier, 2002). That is, dual route models provide a lexical, memory-based route for retrieving spellings of known words and a sublexical or assembled route that relies on P-G mappings and is responsible for spelling unknown words or nonwords. Since most homophones are commonly-known words, their spellings should be retrieved through the lexical route (e.g., Delattre, Bonin, & Barry, 2006). Furthermore, the lexical route generally dominates over the nonlexical one for normal adult spellers (e.g., Folk & Rapp, 2004; Rapp et al., 2002). However, because homophones have two or more spellings for a
shared phonological representation, sublexical activation of P-G mappings might occur when retrieving the appropriate graphemes for each homophone.

Another possibility is that the two routes interact when selecting grapheme candidates, a proposal that has empirical support from lexical priming effects on nonword spelling (e.g., Barry & Seymour, 1988; Campbell, 1983; Folk & Rapp, 2004). The lexical priming effect occurs when a nonword such as /prit/ is spelled as *preet* after hearing *sweet* but as *preat* after hearing *treat* (Campbell, 1983). These findings are offered as evidence that the sublexical process required for spelling nonwords is influenced by the lexical process: the prime word temporarily increases the weighting of the target spelling, therefore increasing the likelihood that the target spelling will be selected upon presentation of the nonword (Barry & Seymour, 1988). In order to more precisely specify the source of this priming, Folk and Rapp (2004) used primes that overlapped in phonology and orthography, or only in orthography, with nonwords. Although priming was found in both conditions, the greater amount of priming following primes that overlapped in phonology and orthography suggests that lexical priming of nonwords relies on preactivation of specific P-G mappings, with a smaller amount of preactivation going to target graphemes. In contrast to nonwords, recent research on lexical priming of real words showed that primes sharing orthography and phonology (e.g., *chaplain*) or only orthography (e.g., *ordain*) with a target word equivalently increased the likelihood of correctly spelling low-frequency, irregular targets (e.g., *porcelain*; Abrams, Trunk, & White, in press). These results suggest the sublexical route can influence spellings of known words that were thought to be retrieved via the lexical route, and that lexical priming of words is driven by orthographic overlap, not P-G mappings.

Whereas previous research investigated priming effects on nonword or low-frequency word spelling, the present research investigated lexical priming of homophones. Homophones
were chosen because they are common words that people generally know how to spell. Furthermore, homophone substitutions are not traditional spelling errors because even though they are used inappropriately within a specific context, they still represent correctly-spelled words. The present experiments investigated the sources of these unique errors by determining whether errors increased following phonological or orthographic primes, and whether errors were more likely for primary (i.e., dominant) or secondary (i.e., subordinate) meanings of homophones. Using laboratory-induced speech errors as a model, we created an error induction paradigm to examine the homophone substitutions that occur when a contextually inappropriate word (*beach*) is written in place of its contextually appropriate homophone (*beech*), e.g., “The teacher enjoys having the beach tree in his yard.” By embedding both primes (*teacher*) and targets (*beech*) in sentence contexts, this paradigm is also the first that we know of to study lexical priming within sentence contexts, which more closely resembles everyday writing and spelling tasks than previous studies using individual words or nonwords.

**Experiment 1**

This experiment investigated lexical priming of dominant and subordinate homophones using primes that overlapped in both phonology and orthography with homophone targets. According to dual route models, homophones should not be influenced by lexical primes because their spellings are retrieved from memory (i.e., the lexical route) and are not influenced by preactivation of P-G mappings or target graphemes (e.g., Delattre et al., 2006). However, as argued by Rapp et al. (2002), integration of the lexical and sublexical routes could occur when both routes simultaneously activate grapheme candidates, and the route with the strongest activation is selected. If this is the case, only dominant homophones should be easily activated through the lexical route, leading to few errors. In contrast, sublexical activation of words may
Homophone Substitution Errors

occur if lexical activation is weaker than sublexical activation, a situation that might arise for the less common (i.e., subordinate) of two homophones. In this case, preactivation of P-G mappings (via exposure to the prime word) should increase the likelihood that those mappings would again be selected when spelling the homophone target. If the prime word shares graphemes with the dominant homophone, then the dominant homophone is likely to be produced, i.e., substituted for the subordinate homophone. Thus, it was expected that homophone dominance would interact with prime condition: Dominant homophones should not be influenced by primes because they are retrieved via the lexical route, whereas subordinate homophones should be influenced by primes because, like nonwords, they are more likely to be retrieved via the sublexical route.

Method

Participants. Fifty-two students from the College of Charleston participated in this experiment. They ranged in age from 18-26 ($M = 19.0, SD = 1.4$), consisted of 39 females and 13 males, and were given partial course credit or extra credit for participating.

Materials

Sentence Writing Task. Ninety-six total homophones (48 homophone pairs) were chosen, and each homophone within a pair was categorized as either dominant or subordinate based on normative estimates (White & Abrams, 2004). Additionally, in order to have consistency between dominance and frequency measures, we chose homophones where the dominant member was also higher in frequency than the subordinate member, consistent with the frequency norms developed by Zeno, Ivens, Millard, and Duvvuri (1995). Overall, dominant homophones had a mean frequency of 290.00 ($SD = 1138.39$) whereas subordinate homophones had a mean frequency of 8.13 ($SD = 17.72$).
Two sentences were created for each of the 96 homophones. Half of the sentences were written so that the dominant homophone was contextually appropriate (i.e., could accurately complete the sentence), and half of the sentences were written to contextually support the subordinate homophone (see Table 1). The sentence context that preceded the homophone’s location in each sentence remained neutral so that the homophone could not be predicted from its preceding context. Rather, the disambiguating context succeeded the homophone and clarified which homophone was appropriate for each sentence. Neutral preceding contexts were pilot tested using a modified cloze procedure: 78 participants were presented with sentences up to but not including the homophones and asked to complete the blank with an appropriate word or phrase. Any sentences that were completed with the target homophone by 10% or more of the pilot participants were revised.

(Table 1 about here)

To measure the influence of primes on homophone spelling, each of the 96 sentences described above were assigned a prime and a corresponding control (i.e., unrelated) word. Thus, two versions of each sentence were created, one that included a prime that overlapped in phonology and orthography with the contextually inappropriate homophone and the other that included a control word that did not overlap in orthography or phonology with either homophone (see Table 1). In most cases, the prime (e.g., teacher) overlapped in the critical differentiating part of the homophone (e.g., beech/beach), but typically also included other parts of the target (e.g., -ch). In the control sentences, the prime word was replaced with an unrelated word that was similar in the length ($M = 5.23, SD = 1.50$ for prime words; $M = 5.64, SD = 1.96$ for control words) and meaning to the prime. The prime/control word was positioned between four and eight words ($M = 6.0, SD = 1.37$) before the homophone in each sentence. This priming manipulation
resulted in 48 sentences priming dominant homophones, 48 control sentences with dominant homophones as targets, 48 sentences priming subordinate homophones, and 48 control sentences with subordinate homophones as targets. Sentences were then counterbalanced so that a participant encountered 12 sentences of each prime/dominance type.

Sentences were audiotaped by a female experimenter and played over Bose® MediaMate® computer speakers via a program written in Visual Basic 5.0.

**Homophone Recognition Test.** A homophone recognition test was constructed to assess participants’ knowledge of the spelling and meaning of individual homophones. For each of the 96 homophones, one sentence was constructed so that it could only be completed with the contextually appropriate homophone (e.g., *The surfers arrived early at the (beach), or (Beech) nuts are a good addition to any salad.*). Sentences were visually presented with a blank space in place of each homophone. Both members of the homophone pair followed each sentence so that participants could circle the correct homophone to fill in the blank. Counterbalancing of the sentences and homophone order resulted in four versions of the recognition test.

**Procedure.** Participants were tested in groups of one to eight per session. Participants were given the sentence writing task first, were told that they would hear 48 auditorily-presented sentences, and were asked to write down verbatim what they heard. A tone was played to indicate the upcoming sentence, followed by a one-second pause, and then the sentence was presented. Participants were told to start writing upon hearing the first word in each sentence. After 30 seconds, another tone was played to indicate that participants should cease writing. At that time, the sentence was repeated if any participant needed to hear it again, and participants were given an additional 15 seconds to write. They were instructed that they could print or write in cursive but that it was important to make their writing legible. They were asked not to mark
through any written responses, and if they wished to make a change, to write the new word(s) in
the space above each sentence. Following administration of the sentence writing task,
participants were given the homophone recognition test to complete at their own pace.

Results and Discussion

The homophone recognition test was used to determine any homophones with which
participants were unfamiliar or of which they did not know the correct spelling. Target
homophones that were not correctly chosen in the recognition test by each participant were
excluded from analyses ($M = 1.06$ homophone pairs per participant, or 2% of the 48 homophone
pairs). If a target was misheard as a different word (3.1%) or not produced (1.4%), it was also
excluded from analyses. Additionally, if a prime was misheard as a different word (1.4%), not
produced (0.8%), or misspelled (1.0%), the target that corresponded to that prime was excluded
from analyses.

The means and standard errors for the percent of homophone substitution errors (e.g.,
writing *beach* in place of *beech*) for prime condition and dominance are reported in Table 2. A 2
(prime condition: primed, control) x 2 (dominance: dominant, subordinate) repeated-measures
analysis of variance (ANOVA) was performed on the mean proportion of homophone
substitution errors by participants ($F_1$) and by items ($F_2$)$^1$. A main effect of prime condition ($F_1$
$1, 51) = 4.87, MSE = .01, p < .032; F_2 (1, 92) = 5.34, MSE = .01, p < .023), revealed that more
errors were made following prime words (11.5%) than control words (7.9%). A main effect of
dominance, $F_1 (1, 51) = 111.79, MSE = .01, p < .001; F_2 (1, 92) = 23.15, MSE = .04, p < .001,
revealed that more errors occurred for subordinate (17.2%) than for dominant (2.2%) homophones. The Prime Condition x Dominance interaction was not significant, $F_1 < 1, F_2 < 1$.

(Table 2 about here)
Experiment 1 demonstrated that both homophone dominance and lexical priming influenced homophone substitution errors. Dominance influenced substitution errors such that dominant homophones were more likely to substitute for subordinate homophones than the reverse. These findings are consistent with research that shows more speech errors and spelling errors for low-frequency than high-frequency words (Dell, 1986; Margolin & Abrams, 2007; Stemberger & MacWhinney, 1986). With regards to lexical priming, more substitutions were found when the contextually appropriate homophone (beech) was preceded by a word (teacher) spelled similarly to the contextually inappropriate homophone (beach). This priming effect was found for both dominant and subordinate homophones, contrary to our predictions. Finding lexical priming of both dominant and subordinate homophones suggests that dominant homophones are not solely retrieved via the lexical route, but are subject to influence from the assembled, sublexical route. A homophone’s spelling can be preactivated by an orthographic unit [EA] through a prime (teacher), increasing the likelihood that that orthographic unit will be selected upon later presentation of the same phonology. One explanation for the sublexical influence on dominant homophone retrieval is that some dominant homophones are quite low in frequency (e.g., ale has a frequency of 1; Zeno et al., 1995), despite being the more common meaning of the homophone pair (White & Abrams, 2004). Thus, dominant homophones that are low in frequency might rely on sublexical processes in order to assemble their spellings, leaving them open to lexical priming. However, the overall low rate of errors for dominant homophones suggests that they are usually retrieved from memory via the lexical route.

Experiment 2

Experiment 2 investigated whether the lexical priming found in Experiment 1 was due to an overlap in phonology and orthography between the prime and target, or simply to
orthographic overlap. Larger lexical priming effects on nonwords (/n tS/ spelled as NOUCH) occur when primes overlap in phonology and orthography (e.g., couch) with the target, compared to orthographic-only primes (e.g., touch) (Folk & Rapp, 2004). In contrast, lexical priming of specific irregular spellings within low-frequency words is dependent on orthographic, but not phonological, overlap between primes and targets (Abrams et al., in press). If homophone priming behaves similar to nonword priming and preactivates specific P-G mappings, primes that overlap in phonology and orthography should produce greater substitutions than primes that overlap in orthography only. However, if homophone spellings can be preactivated by presenting specific graphemes independent of phonology, then there should be no differences between the two prime conditions.

Method

Participants. This experiment included 50 students from the College of Charleston who were compensated through course or extra credit. They ranged in age from 18-32 (\(M = 20.3, SD = 2.7\)) and included 33 females and 17 males.

Materials

Sentence Writing Task. The sentence writing task was similar to Experiment 1, except only 36 of the 48 subordinate homophones used in Experiment 1 were included because of difficulties finding primes for both the phonological and the orthographic conditions. Only subordinate homophones were used because of the infrequent occurrence of errors on dominant homophones in Experiment 1. Sentences for the 36 homophones were written so that the subordinate homophone was contextually appropriate. One word was used for each prime condition: For the phonological condition, primes shared both phonology and orthography (e.g., teacher) with the dominant homophone (i.e., contextually inappropriate homophone, beach). For
the orthographic condition, primes only overlapped in orthography with the dominant homophone (e.g., headmaster). For the control condition, the primes were replaced with a word that did not overlap in orthography or phonology with the target but maintained the contextual meaning of the sentence (e.g., lawyer).

Similar to Experiment 1, sentences were audiotaped by a female experimenter and played over Bose® MediaMate® computer speakers via a program written in Visual Basic 5.0.

**Homophone Recognition Test.** Similar to Experiment 1, a homophone recognition test was constructed and counterbalanced for the 36 homophones used in Experiment 2.

**Procedure.** The procedure was identical to Experiment 1.

Results and Discussion

Target homophones that were not correctly chosen in the homophone recognition test by each participant were excluded from analyses (M = 0.88 homophone pairs per participant, or 2.4% of the 36 homophone pairs). If a target was misheard as a different word (6.2%) or not produced (3.6%), it was also excluded from analyses. Additionally, if a prime was misheard as a different word (1.0%), not produced (1.04%), or misspelled (2.7%), the target that corresponded to that prime was excluded from analyses.

The means and standard errors for the percent of homophone substitution errors for each prime condition are reported in Table 3. An ANOVA with prime condition (phonological, orthographic, control) on the mean proportion of homophone substitution errors was significant, $F_1 (2, 98) = 5.76, MSE = .01, p < .004; F_2 (1, 35) = 3.25, MSE = .01, p < .045$. Follow-up analyses showed more errors following phonological primes than control words, $p_1 < .001, p_2 < .026$, and following orthographic primes than control words, $p_1 < .025, p_2 < .060$. There were no differences in errors following phonological and orthographic primes, $p_1 > .381, p_2 > .468$. 
(Table 3 about here)

Experiment 2 showed that phonological overlap between the prime and target was not necessary to increase homophone substitution errors, which occurred when primes only shared orthography with targets. Finding priming in both conditions supports similar findings using nonwords (Folk & Rapp, 2004) and low-frequency words (Abrams et al., in press). However, in contrast to Folk and Rapp (2004), we did not find a significantly larger priming effect for words overlapping in phonology and orthography than words overlapping in orthography only, suggesting that the preactivation of P-G mappings is not solely responsible for spelling via the sublexical route, consistent with Abrams et al. (in press).

General Discussion

These experiments showed that both homophone dominance and lexical priming influenced written homophone substitution errors. These findings support and extend recent research with nonwords that the lexical and sublexical routes can interact to influence spellings of words (e.g., Rapp et al., 2002). Our findings suggest that the spellings of higher-frequency words (or dominant homophones in this case) receive strong activation via the lexical route and therefore rarely result in errors. In contrast, a situation arises when trying to spell subordinate homophones (and possibly low-frequency dominant homophones) where sublexical activation is stronger than lexical activation, resulting in grapheme retrieval via the sublexical route. When graphemes that correspond to a dominant homophone have just been used (i.e., by spelling prime words), the result is a dominant homophone substituting for its subordinate partner.

The lexical priming effect complements similar findings where recent activation of a specific orthographic component increased correct spelling of that same component in a difficult-to-spell word (Abrams et al., in press) and extend these findings to show that exposure
to orthography can also hurt spelling by encouraging retrieval of a wrong word. Furthermore, finding that primes only needed to overlap in orthography with targets to increase errors indicates that retrieval of real word spellings via the sublexical route does not always invoke phonology, contrary to retrieval of nonword spellings (Folk & Rapp, 2004). One explanation for these priming differences between words and nonwords is that primes were heard and then produced in the present study, but were only perceived in Folk and Rapp’s study, suggesting the nature of writing the prime might over-emphasize reliance on orthography. Alternatively, the discrepant results might be explained by existing, well-established representations in memory for real words, compared to nonexistent representations for nonwords. Reliance on phonology might be necessary when spelling a word that does not exist, similar to the importance of phonology when children are learning to spell (e.g., Treiman, 2003). In contrast, assembling spellings of known words allows phonology to be bypassed, similar to using the lexical route to retrieve whole words from memory.

Finding that assembled spellings can rely only on orthography has several broader implications for research on spelling as well as language production more generally. First, these findings suggest that retrieval via an assembled orthographic route not only influences correct spelling within words (as in Abrams et al., in press), but can also be applied to homophones which are always correctly spelled, but contextually inappropriate. That is, the substitution errors demonstrated here offer a situation where assembled orthography (even if it is just of an orthographic unit) results in selection of an entire, inappropriate word. Furthermore, demonstrating reliance on an assembled orthographic route using homophones shows that specific spellings can be primed and that these spellings dominate over other information available for word retrieval, such as the constraints given by context (i.e., a specific meaning).
Parallel findings have been demonstrated in other language production domains, such as resolution of tip-of-the-tongue states, where phonological cues increased resolution more than semantic cues (e.g., Meyer & Bock, 1992).

In conclusion, these two experiments have shown that homophone substitution errors, which involve correct spellings out of context, are subject to similar influences as traditional spelling errors involving incorrect spellings. Future research might investigate whether substitution errors for subordinate homophones are differentially affected by low- or high-probability spelling patterns for a shared phonology, or whether semantics interacts with orthographic retrieval to influence substitutions via sentence context.
References


Author Notes

This research was supported by a Sigma Xi Grant-in-Aid of Research, awarded to Sarah Zoller, and by a College of Charleston Summer Undergraduate Research Fellowship, awarded to Sarah Zoller and Katherine White. Portions of this research were reported at the 11th biennial Cognitive Aging Conference in April 2006 and the 48th annual meeting of the Psychonomic Society in November 2007. We thank Christopher Barry for his helpful comments during a discussion of this paper. Correspondence concerning this article should be sent to Katherine White, Department of Psychology, College of Charleston, 57 Coming Street, Charleston, SC 29424. Email: whitek@cofc.edu.
Footnotes

1. Due to an error in the computer program, two of the homophone sentences (the primed sentence for the dominant homophone *bail* and the control sentence for the dominant homophone *idle*) were never presented. Thus, 94 instead of 96 stimuli were used in the item analyses.
Table 1

*Example Primed and Control Sentences for Dominant (beach) and Subordinate (beech)*

*Homophones used in Experiment 1.*

<table>
<thead>
<tr>
<th>Prime Type</th>
<th>beach (dominant)</th>
<th>beech (subordinate)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primed</td>
<td>After presenting her <em>speech</em> on animal rights, Sue went to the <em>beach</em> to relax.</td>
<td>The <em>teacher</em> was most proud of the <em>beech</em> tree in his garden.</td>
</tr>
<tr>
<td>Control</td>
<td>After presenting her <em>talk</em> on animal rights, Sue went to the <em>beach</em> to relax.</td>
<td>The <em>lawyer</em> was most proud of the <em>beech</em> tree in his garden.</td>
</tr>
</tbody>
</table>

Note. Primed and control words are underlined, and target homophones are italicized.
Table 2

*Mean Percent of Homophone Substitution Errors in Experiment 1 by Dominance and Prime Condition.*

<table>
<thead>
<tr>
<th>Prime Condition</th>
<th>Dominant</th>
<th>Subordinate</th>
<th>Dominant</th>
<th>Subordinate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>(SE)</td>
<td>% Priming</td>
<td></td>
</tr>
<tr>
<td>Primed</td>
<td>3.3</td>
<td>(0.8)</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>19.6</td>
<td>(2.7)</td>
<td>4.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.1</td>
<td>(0.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>14.7</td>
<td>(1.6)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3

*Mean Percent of Homophone Substitution Errors in Experiment 2 by Prime Condition.*

<table>
<thead>
<tr>
<th>Prime Condition</th>
<th>Phonological</th>
<th>Orthographic</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>M</strong></td>
<td>23.3</td>
<td>21.0</td>
<td>15.6</td>
</tr>
<tr>
<td><em>(SE)</em></td>
<td>(2.8)</td>
<td>(2.4)</td>
<td>(1.6)</td>
</tr>
<tr>
<td>% Priming</td>
<td>7.7</td>
<td>5.4</td>
<td></td>
</tr>
</tbody>
</table>