Winter 1-2021

# Spelling and Reading Novel Homophones: Testing the Value of Lexical Distinctiveness

Jayde Homer

WASHINGTON UNIVERSITY IN ST. LOUIS

Department of Brain and Psychological Sciences

# Spelling and Reading Novel Homophones:

# Testing the Value of Lexical Distinctiveness

by
Jayde Elizabeth Homer

A thesis presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Master of Arts

January 2021
St. Louis, Missouri

# Table of Contents

# List of Tables

# Acknowledgments

I would like to express my great appreciation to Dr. Treiman for her valuable contributions and support during the course of this thesis. I gratefully acknowledge Drs. Butler and Van Engen for their support. The completion of this thesis would not have been possible without: Dr. Kessler's guidance and expertise, the Reading and Language Lab research assistants' time, and the many undergraduates who participated in these experiments. A special thanks goes to the Washington University Department of Psychological and Brain Sciences for awarding me a summer grant to financially support a portion of my data collection.

I wish to thank my parents, grandparents, and partner for their endless support and encouragement. And finally, I cannot begin to express my thanks to my cats, who never wavered in their love and support while a pandemic plagued the world and I completed this thesis from home.

Jayde Elizabeth Homer

*Washington University in St. Louis*

*January 2021*

ABSTRACT OF THE THESIS

Spelling and Reading Novel Homophones:

Testing the Value of Lexical Distinctiveness

by

Jayde Elizabeth Homer

Master of Arts in Brain and Psychological Sciences

Washington University in St. Louis, 2020

Professor Rebecca Treiman, Chair

Lexical distinctiveness, according to which a written form represents one and only one morpheme, is a feature of some writing systems. For example, ‹bear› and ‹bare› are spelled differently in English. In two experiments, we asked whether readers and spellers of English benefit from distinctive spellings of homophones. In Experiment 1, university students listened to 40 passages, each containing a novel homophone (e.g., /kel/ used to mean a gossip-lover). In Experiment 2, participants read the passages. Half of the novel homophones were homographic (e.g., ‹kale›), and half were heterographic (e.g., ‹kail›). In both experiments, participants answered questions about the novel word either by choosing between two spelling options (e.g, ‹kale› vs ‹kail›) or producing their own spelling. We also asked participants about whether new words in a language should have distinctive spellings. The majority of participants in both experiments expressed a preference for distinctive spellings. Experiment 1 participants chose heterographic spellings more frequently than homographic ones; however, they produced more homographic than heterographic spellings. Experiment 2 participants recalled heterographic and homographic homophones from the passages at equal rates, suggesting they did not benefit from distinct spellings of novel homophones. These findings lead us to question whether lexical

distinctiveness is an essential feature of the English writing system, as some linguists have theorized. We explore various factors that work against lexical distinctiveness, including the distribution of heterographic homophones in the English language, the challenge of generating novel spellings, and the unclear division between polysemy and homophony.

# 1. Introduction

Lexical ambiguity is ubiquitous in natural language, often leaving the meaning of a word open to a listener's interpretation. Most words in English have multiple interpretations (Rodd et al., 2002). The two forms of lexical ambiguity are syntactic and semantic (Small et al., 1988). Syntactic lexical ambiguity is a category issue, arising when a word's part of speech is unclear. Semantic lexical ambiguity is a meaning issue and is the focus of the present study. One form of semantic lexical ambiguity, homonymy, occurs when a word has multiple unrelated definitions. Consider the two uses of *bat*[1] in "After hitting a grand slam, the player tossed the *bat* towards the dugout" and "Molly had to get a series of rabies vaccines after finding a *bat* in her room." The intended meaning of *bat* in both phrases is clear, and the examples illustrate the two unrelated definitions the word carries. Yet when context is insufficient, as in "Molly found a *bat* in her room," the meaning of *bat* is unclear, and the listener is likely to misinterpret or be confused. Another form of semantic lexical ambiguity is polysemy. This occurs when a word's multiple meanings are related; polysemous words often share a root or origin. For example, the uses of *book* in "Hand me the *book* so I can read" and "Can you spare a match from your *book*?" are similar but do not have the same literal meaning. Lexical ambiguity is a difficult problem that language users must work to resolve or avoid so that they can maintain clear communication.

An important question is whether a language's writing system can reduce semantic lexical ambiguity. In the previous examples, the words share a phonological and orthographic form (i.e., *bat* and *bat* sound and are spelled the same). However, not all words that sound the same are spelled the same. In English, and some other languages with alphabetic writing

[1] Note on symbols: ‹› are used for orthographic representations, italics are used for lexical representations, and / / are used for phonetic representations.

systems, there are multiple ways to spell a single sound. For example, the sound /i/ can be spelled as ‹ea› or ‹ee›. The phrase “There is a /lik/ in the sink” has quite a different meaning depending on whether /lik/ is spelled ‹leak› or ‹leek›. Because a listener only has information provided by the word’s sound form and context, the sentence is ambiguous. However, in writing, “There is a *leek* in the sink” is not ambiguous. The reader has an additional cue that the listener does not have: spelling. In this case, the spelling disambiguates the intended meaning. Some writing systems, including English, can capitalize on their ability to represent one sound with multiple letters to allow different spellings for words that happen to sound the same and thereby reduce lexical ambiguity.

Some theories about how writing systems should work suggest that languages should differentiate spellings of words that sound the same (e.g., Venezky, 1999). As a basic principle, lexical distinctiveness requires that each word form be visually distinguished from other word forms (Berg & Aronoff, 2020; Rutkowska & Rossler, 2012; Ryan, 2016; Kessler & Treiman, 2015). For example, the combination of letters ‹onion› represents *onion* and never another word, like *pepper*. Lexical distinctiveness is particularly valuable in reducing semantic lexical ambiguity when two words sound the same. Distinct spellings (e.g., ‹leak› and ‹leek›) ensure that a reader can distinguish between the two words and interpret the intended meaning. A writing system completely lacking in lexical distinctiveness would spell all words the same, while a writing system that adheres to lexical distinctiveness would have no words with shared spellings and could avoid semantic lexical ambiguity. When a writing system uses differentiates two words that sound the same with distinct spellings, the result is a pair of heterographic homophones (e.g., *leak*/*leek, whale*/*wail, ate*/*eight*). Heterographic homophones help to reduce semantic lexical ambiguity that is present in the spoken word form. Although heterographic

homophones follow a principle of lexical distinctiveness and encourage clarity, many writing systems, including English, do not perfectly adhere to lexical distinctiveness. These violations result in homographic homophones[2], words that sound and are spelled the same (e.g., *duck* or *bat*). Without sufficient context, homographic homophones maintain semantic lexical ambiguity and make it difficult for a reader to interpret the intended meaning. A writing system that adheres to lexical distinctiveness would not have homographic homophones or the semantic lexical ambiguity they cause.

Those who speak, read, and write in English are aware that heterographic homophones reduce ambiguity and confusion. This awareness, some behavioral evidence suggests, leads to a preference for distinctiveness. Baker (1980) tested this awareness by examining whether English spellers ($n$ = 11), playing the role of language reformers, would treat lexical distinctiveness as a desirable feature of the language. Participants rated the rationality of existing homophone spellings on a scale of 1 to 5 (5 = "perfectly rational"). For example, participants saw words like *seam*, rated the rationality of the spelling, and then provided a reformed spelling for any word rated less than perfectly rational. Overall, participants showed a significant tendency to avoid homographic homophones. For example, participants who rated *seam* as less than perfectly rational often produced a novel spelling like *seme* instead of the existing homophone counterpart, *seem.* Baker suggested that the participants were expressing a preference for heterography because they were aware of the ambiguity of homographic spellings, and they avoided homographic spellings when given the opportunity to provide a new spelling for an existing word.

[2] Note: the term homographic homophone is used here interchangeably with what is often referred to as a *homonym*.

People seem to prefer distinctive spellings of homophones (Baker, 1980) and theories of writing systems suggest that lexical distinctiveness should be maintained, yet English has many homographic homophones. To better understand why language users endorse a feature that is not reliably present in their written language, we can examine the behavior of adult spellers learning new homophones. Baker's results suggest that, when given the opportunity to spell a familiar sounding word that has a novel meaning, people should provide a distinctive spelling.

Treiman, Seidenberg, and Kessler (2015) sought to assess how linguistic and performance factors influence spelling and replicate the original metalinguistic findings of Baker (1980) that people prefer heterography. In addition to explicitly asking about heterography preferences, as Baker did, the researchers asked whether university students would spell novel homophones differently from the existing words. In the first of three experiments, participants heard a novel definition for a familiar phonological form, for example that /wˈɪntɚ/ is a person who is eager to learn the latest news and gossip. Participants were then asked to spell the word. There are several potential ways to spell /wˈɪntɚ/: like the corresponding familiar word ‹winter› or with a novel spelling such as ‹whinter› or ‹winnter›. If participants prefer distinctiveness, then they should produce a novel, heterographic spelling more often than they reuse an established spelling. However, 66% of responses provided for novel homophones like /wˈɪntɚ/, /ˈkrɪkət/, and /ˈsɛvən/ were established spellings like ‹winter›, ‹cricket›, and ‹seven›. Contrary to Baker's findings and the theoretical property of lexical distinctiveness, the majority of participant productions were homographic. What pressures might have led to these counterintuitive results?

In order to address this question in a second experiment, the researchers implemented a metalinguistic awareness manipulation by informing participants that each word could be spelled in more than one way. Participants were encouraged to think about the best way to spell each

word. The researchers expected that this manipulation might encourage participants to think about their preference for heterography and follow through with it by producing a distinct spelling. At the end of the experiment, participants were explicitly asked whether it is better for a new word entering English to be spelled like an existing word that sounds the same, or differently. Participants were more likely to state that a novel homophone should have a novel spelling than to produce a novel spelling in the experiment. Neither the metalinguistic awareness manipulation nor participants' stated preference had an effect on spelling production. The results of these two experiments confirm that people are aware of the ambiguity that results from a lack of distinctiveness yet fail to use that knowledge when spelling novel homophones.

To identify the source of this discrepancy, Treiman et al. (2015) conducted a final experiment with a manipulation of spelling condition. Participants were again auditorily presented with a familiar phonological form (e.g., /ˈwɪntɚ/) with a novel definition read aloud by an experimenter. Half of the participants were assigned to a condition in which they were asked to choose between two possible spellings for each of 13 novel words: one established and one novel. For /ˈwɪntɚ/, for example, participants chose between the established spelling ‹winter› and the novel spelling ‹whinter›. The other participants were assigned to a production condition like that of the previous two experiments in which they were asked to spell the novel words. The results of the production condition in Experiment 3 did not differ from the first two experiments, with 60% established spellings produced for homophone items. However, participants in the choice condition chose the established spellings only 27% of the time. When given a choice, participants exhibited a preference for distinctiveness by choosing mostly heterographic spellings. When producing a spelling, participants often created semantic lexical ambiguity by providing homographic spellings.

Why do participants reuse spellings when they are aware of the ambiguity that homographic homophones cause? The explanation might lie in the ease of reusing an established spelling. Transcribing a word is more cognitively demanding than seeing and choosing a word. It requires listening, holding the sound in memory, translating the sound to a spelling, and producing the spelling. When participants hear a familiar phonological form, the established spelling is activated in their mental lexicons. In order to abide by their stated preference and spell the familiar phonological form with a novel spelling, participants must overcome interference from prior knowledge, such as the established spelling or meaning (Fang et al., 2017). Suppressing competing knowledge could be made easier by seeing multiple plausible spellings. When seeing two spellings, a participant simply recognizes and chooses the word that best matches their intuition about how new words should be spelled. When not given spelling choices, it is more efficient to access a spelling that is already stored, reducing cognitive effort. The results of Treiman et al. (2015) suggest that, in this case, participants sacrifice unambiguous language for ease when spelling new words.

Before accepting the conclusion that people sacrifice distinctive, unambiguous spellings for ease, it is necessary to consider several limitations of Treiman et al. (2015) and replicate the results. First, their final experiment included only 13 novel homophones. Second, potential variability was introduced by having an experimenter read items and definitions to participants. It is likely that the experimenter differed in reading rate, volume, pronunciation of the novel word, patterns of inflection, stutters or mistakes, prosody, and body language with each presentation of an item and definition. This variance may have influenced participants' responses and be unaccounted for in the data. Third, the learning paradigm (i.e., item and definition) does not reflect a natural vocabulary learning setting. Rarely is a new word learned outside of a

classroom by hearing it with an explicit definition. People learn new vocabulary words incidentally from conversations, podcasts, radio, audiobooks, television, and movies. In doing so, they must assume each unfamiliar words' meanings from context. It is socially beneficial to be able to infer meanings of words and disambiguate on the fly given context.

To overcome these limitations and test the generalizability of Treiman et al. (2015), we designed two experiments with the aim of replicating and extending the finding that people prefer heterographic homophones, yet often do not produce them spontaneously. In Experiment 1, participants listened to short passages to learn novel words in context. In Experiment 2, we had participants read the same passages to learn the spelling of a novel word. In both experiments, we addressed the limitations of Treiman et al. that were described above. We expanded the number of experimental items from 13 to 80. Each item shared a phonological form with a familiar English word. To eliminate variability introduced by the experimenter, participants completed the study in a computer program; we controlled the speed, volume, and clarity of the stimuli. To make the learning paradigm more natural, we used an incidental vocabulary acquisition task. As opposed to presenting novel homophones with formulated definitions, we embedded each novel homophone in a fictional passage.

In Experiment 1, we asked whether participants would spell familiar-sounding novel words differently depending on how the spelling was elicited. Specifically, if the results of Treiman et al. (2015) are replicable, we predict that participants in the choice condition will choose a heterographic spelling more often than a homographic spelling, while participants in the production condition will produce a homographic spelling more often than a heterographic spelling. For example, we expect participants who hear /naʊn/ to choose the heterographic spelling ‹nown› if they are in the choice condition, while participants in the production condition

are expected to produce the homographic spelling ‹noun› more often than a heterographic spelling. This pattern of results would support the argument that reusing established forms is easier than creating a novel spelling, even though reused spellings cause semantic lexical ambiguity.

In Experiment 2, we extended Treiman et al. (2015) and the present Experiment 1 to incidental vocabulary acquisition in silent reading. We asked whether condition (choice or production) and homophone spelling (heterographic or homographic) affects recall of the spellings of novel homophones. That is, does people's memory for new words differ depending on whether a spelling is heterographic or homographic to a familiar English word? To test this, participants read the short passages from Experiment 1 containing a novel word that sounds like a familiar English word (e.g., /naʊn/). Half of the novel words had homographic spellings (e.g., ‹noun›), while the other half had heterographic spellings (e.g., ‹nown›). After participants read each passage, we tested their memory of the novel homophone by asking a question eliciting its spelling. Half of the participants were given a blank line and produced a spelling, while half were given two options to choose from, like ‹noun› and ‹nown›. People can learn new words in silent reading, and they do this well even with homophones (Brushnighan et al., 2014; Folk, 1999; Van Orden et al., 1988). However, previous studies have not directly compared the learning of heterographic and homographic spellings of novel homophones.

# 2. Experiment 1

## 2.1 Method

### 2.1.1 Participants

Participants were 39 individuals from the Washington University in Saint Louis undergraduate subject pool. They received course credit for their participation. All participants had normal or corrected-to-normal vision and normal hearing. Three additional people took part in the experiment, but their data were not analyzed because they were not native speakers of English. The mean age of the participants whose data were used was 20 years (range 18–25).

### 2.1.2 Stimuli

A self-paced 40-trial spelling task was programmed in PsychoPy3 (Peirce et al., 2019) and took approximately 25 minutes for participants to complete. Each trial consisted of three audio recordings: a passage containing a novel word, a comprehension question, and a vocabulary question. The vocabulary questions were designed to elicit a spelling of the novel homophone presented in the passage.

To construct the novel homophones for the study, we created 80 English monosyllabic phonological forms that have at least two plausible spellings. Table 2.1 shows examples of the phonological forms and their spellings (see the Appendix for the list of all homophone items). All of the phonological forms mapped onto common English words, such as /naʊn/. Each phonological form had one spelling that was conventional, in this case ‹noun›. We refer to this as the established spelling. This phonological form may also be plausibly spelled as ‹nown›. We refer to this as a novel spelling. The phoneme whose spelling differs between the established spelling and the novel spelling, in this case /aʊ/, is called the critical phoneme. Spellings of a given critical phoneme may appear in English with varying frequencies; that is, it may be much

more common to see one spelling of /aʊ/ than another. To control for this, we designed the stimuli such that, for half of the items with each critical phoneme, one spelling of this phoneme appeared in the established spelling. For the other half of the items with each critical phoneme, the other spelling of the critical phoneme appeared in the established spelling. For example, a participant hears /klaʊn/, which they know as ‹clown›, with ‹ow›. Elsewhere in the experiment, they hear /naʊn/, which has the established spelling ‹noun›, with ‹ou›. For each item, the established and novel spellings were used as the answer choices in the choice condition.

**Table 2.1** Sample Stimuli for Experiment 1 and Experiment 2

| Phonological form | Critical phoneme | Established spelling | Novel spelling |
|---|---|---|---|
| /naʊn/ | aʊ | ‹noun› | ‹noun› |
| /klaʊn/ | aʊ | ‹clown› | ‹cloun› |
| /ʤel/ | e | ‹jail› | ‹jale› |
| /kel/ | e | ‹kale› | ‹kail› |
| /fil/ | i | ‹feel› | ‹feal› |
| /dil/ | i | ‹deal› | ‹deel› |
| /ræθ/ | r | ‹wrath› | ‹rath› |
| /ræg/ | r | ‹rag› | ‹wrag› |

In order to keep the task length around twenty minutes and avoid disengagement, we divided the phonological forms into two 40-item lists. Each list had the same number of items with each critical phoneme (e.g., /aʊ/ is in two phonological forms in both lists, so each participant heard /aʊ/ twice). Participants were randomly assigned to one of the two lists.

We chose the 80 phonological forms based on the results of a screening study with 40 participants from the same population who did not participate in the experiment itself. In the screening study, we tested 258 phonological forms with an established spelling and a novel spelling that we thought participants would be likely to endorse (e.g., /naʊn/ spelled ‹noun› and ‹nown›). We asked participants in the screening study to read aloud a randomized list of established spellings and novel spellings. For example, the list included established spellings such as *food* and novel spellings such as *nown* (for *noun*), *pyth*, *skeight*, *foud*, *cloun*, and *quale*. An experimenter transcribed the participants' pronunciations. After reading the list, participants were asked to circle all the items that they "believe to be correctly spelled, real English words." For a phonological form to be included in the experiment, it had to be the case that (a) the established spelling was pronounced correctly more than 80% of the time, (b) the novel spelling was pronounced as the corresponding real word more than 80% of the time, (c) the established spelling was categorized as a real English word more than 90% of the time, and (d) the novel spelling was categorized as a real English word less than 10% of the time. We constructed the final list of 80 phonological forms from the items that met all of these criteria.

To carry the novel words, we wrote 40 passages with an average length of 37 words (range 23-53). Table 2.2 shows examples of the passages. The novel word was in the first sentence in all passages. The passage was meant to convey the meaning of the novel word. We attempted to ensure the novel word in each passage was semantically unrelated to the passage. For example, because the first passage in Table 2.2 describes a plant, we did not use the items /wid/ or /wit/ because they refer to plants.

**Table 2.2** Sample of Experiment 1 Audio Passages, Questions, and Alternatives in Choice Condition

| Passages | Questions | | Choice condition alternatives | |
|---|---|---|---|---|
| | Vocabulary | Comprehension | Vocabulary | Comprehension |
| Natalie stared at the /naʊn/ soaking up the warm sun on the shelf and tried to remember the last time she watered it. She had never noticed how much the spiky plant resembled a tiny round person with a big tuft of orange hair atop his head. | What type of plant does Natalie have? | Where was the plant sitting? | nown<br>noun | on the table<br>on the shelf |
| Maggie reached up to the lowest branch of the /klaʊn/ and held its tiny yellow berry in her hand. Tears came to her eyes as she reminisced about the days climbing these trees and eating the sweet berries with her sister. | What is the name of the tree that Maggie and her sister used to climb? | What color was the berry Maggie held? | clown<br>cloun | red<br>yellow |

We constructed two *wh*-questions for each passage that were designed to minimize the possibility of guessing the correct answer. For the second passage shown in Table 2.2, for example, the comprehension question asked about the color of the berry. This question could not be answered correctly by using knowledge about the most common colors of berries; it required information from the passage. The vocabulary question for this passage asked participants to provide the name of the tree.

A woman with a General American accent recorded the passages and questions with a mean rate of 208 words per minute. We normalized recording intensities to 70 dB.

We piloted the passages and questions with 20 participants from the same population who did not participate in the experiment itself. We made small edits to passages and questions if either the vocabulary or comprehension questions were answered incorrectly more than 50% of the time. Because some production questions were never answered correctly in the first round of piloting, we again tested the passages and questions with another 4 participants from the same population to confirm that all questions could be answered correctly more than 50% of the time.

### 2.1.3 Procedure

Each participant was randomly assigned to either the choice or production condition. In the choice condition, participants were given a choice between the established spelling and the novel spelling when answering the vocabulary question (e.g., ‹noun› and ‹nown›). Participants were given a choice between the correct answer and a foil when answering the comprehension question. In the production condition, participants were instructed to type their answers for vocabulary and comprehension questions. The order of the passages, the order of the two types of questions for each passage, and the order of the answer choices were randomized for each participant.

An experimenter informed participants that they would hear a series of passages and would answer two questions about each passage. The participant also read instructions on the computer screen before beginning the task. Depending on the condition to which they were assigned, participants were instructed to either select a choice using one of two keyboard keys or to type their response. Each passage and question played once, with no option of repeating it. After answering the two questions for a passage, participants indicated that they were ready to listen to the next passage by pressing a specified key. Participants heard passages and questions over headphones. At the beginning of the experiment, participants familiarized themselves with the format of the computer task by completing two practice trials with passages that did not contain any novel words. The experimenter stayed with the participant during the practice trials and answered any questions about the task. Once any questions were answered, the experimenter left the room.

After finishing the computer task, participants completed an optional background information survey with questions about race, gender, education, and age. Participants then completed a two-item survey, the metalinguistic questionnaire, similar to that of Treiman et al. (2015). The first question of the metalinguistic questionnaire asked, “When a new word enters a language that has a different meaning but the same pronunciation as an existing word, is it better to spell the new word like the existing word?” Participants were asked to explain their reasoning. The second question asked participants to explain why another person might support the opposing view. At the end of the experimental session, the experimenter debriefed participants, informing them that the words they heard were real English words but were given invented meanings for the purpose of the study.

### 2.1.4 Scoring

For the choice condition, a computer program scored responses. For vocabulary questions, there were two response types: established (‹noun› for /naʊn/) or novel (‹nown› for /naʊn/). For comprehension questions, there were also two response types: correct or incorrect.

For the production condition, all responses to the vocabulary and comprehension questions were scored by two judges for accuracy and agreement. If the two judges did not provide the same score for a response, they discussed the difference and came to an agreement. For vocabulary responses, there were four response types. An established spelling response was the English spelling represented by the phonological form, ‹noun› for /naʊn/. A novel spelling response was either the alternative provided in the choice condition, ‹nown› in the example, or a plausible invented form such as ‹knaun› or ‹nowne› that used the sound-to-letter mappings of English as listed in Cummings (1988). Incorrect responses were responses that did not fit in the first two categories. Incorrect responses included, for example, ‹namu› for /naʊn/, which does not map to the phonological form, a synonym or other semantically related word, or the definition of the novel homophone instead of the novel homophone itself. A no-response included a blank response, *???*, or *I don't know*. For comprehension questions, the judges scored responses as correct or incorrect.

## 2.2 Results

A summary of spelling responses is presented in Table 2.3. We found a difference in novel homophone spelling as a function of condition such that novel spellings were chosen more often than they were produced. In the choice condition, participants chose novel spellings more frequently than established spellings, while in the production condition, participants produced more established spellings than novel spellings. Half of the responses produced were categorized as incorrect or no response and thus were excluded in order to directly compare novel spelling

proportions between the two conditions. In the production condition, the proportion of novel spellings was .25.

**Table 2.3** Experiment 1 Mean Proportions (*SD*) of Spelling Response Type as a Function of Condition

| Condition | Established spelling | Novel spelling | Incorrect response | No response |
|---|---|---|---|---|
| Production | .38 (.14) | .13 (.10) | .44 (.19) | .06 (.09) |
| Choice | .28 (.20) | .72 (.20) | - | - |

To test the primary hypothesis of a condition difference in novel homophone spelling, data were analyzed with logistic mixed-effects models using the lme4 package in R (Bates et al., 2015). In this analysis, only responses that were categorized as established or novel were included, excluding the 50% of responses that were categorized as incorrect or no-response in the production condition. The reference level for the fixed effect of condition was choice. Thus, the intercept in the model represented spelling outcomes for choice condition. Random effect terms included intercepts of participant and homophone item; a by-item slope for condition by phonological form showed a significant contribution in a model comparison. There was a main effect of condition on the log odds of using a novel spelling ($\beta = -2.80$, $SE = .39$, $z = -7.16$, $p < .001$, 95% CI [−3.56, −2.03]). This translates to predicted probabilities of .78 in the choice condition and .18 in the production condition for novel spellings. In the choice condition, participants chose novel spellings most often and at above the level of chance ($t(18) = 4.89$, $p < .001$, 95% CI [.64, ∞]).

For comprehension questions, participants more often responded correctly in the choice condition ($M$ = .90, $SD$ = .06) than he production condition ($M$ = .67, $SD$ = .15) with a difference of .23, two-sample $t(23.08) = 6.27$, $p < .001$, 95% CI [.15, .30]. We found no evidence suggesting that any particular passages or homophone items accounted for a disproportionate amount of incorrect responses on comprehension or vocabulary questions.

Table 2.4 presents the mean proportion of novel spellings as a function of metalinguistic questionnaire response and condition. Regardless of condition, most participants said that novel homophones introduced to the language should be spelled differently than an existing word. For example, one participant in the production condition said that different spellings should be used "to make it less confusing in what meaning the word is intended to have." Similarly, a participant in the choice condition who preferred different spellings said, "if the word is spelled differently, there is less chance for confusion about which definition of the word is being used." These responses are representative of most supporters of heterography; most responses used *confusion, distinguish,* and *meaning* and alluded to ambiguity being bad. The few participants who preferred same spellings offered responses like "it avoids misspelling the word and getting confused between various spellings, with context you would never be confused about the definition." Other participants cited simplicity and the ease of remembering one spelling versus two. The few participants who expressed having no preference regarding homophone spelling provided explanations that reflected lack of interest and engagement in the experiment. The proportion of novel spellings was slightly greater for participants who expressed a heterography preference than for participants who expressed a homography preference. However, these results must be interpreted with caution; given the small number of participants who expressed a

homography or a lack of a preference, we could not conduct any statistical tests on the metalinguistic questionnaire data.

**Table 2.4** Number of Participants (*n*) and Mean Proportion of Novel Spelling Responses by Stated Spelling Preference and Condition

| Condition | Stated Spelling Preference for Novel Homophones | | | | | |
|---|---|---|---|---|---|---|
| | Different | | Same | | No preference | |
| | *n* | Mean | *n* | Mean | *n* | Mean |
| Choice | 16 | .74 | 2 | .55 | 1 | .70 |
| Production | 15 | .13 | 2 | .05 | 3 | .17 |

# 2.3 Discussion

In Experiment 1, we aimed to replicate the finding of Treiman et al. (2015) that participants generally spell familiar-sounding novel words as heterographic in a choice condition and homographic in a production condition. In doing so, we accounted for limitations and generalized the effect to incidental vocabulary acquisition. We found that participants chose novel spellings more than established spellings and produced established spellings more than novel spellings. Participants also explicitly endorsed novel spellings as more favorable than established spellings for novel homophones.

Why do we see a higher proportion of established spellings than novel spellings in the production condition of both the present experiment and Treiman et al. (2015) Experiment 3? We may not have seen a reliance on established forms in the choice condition because the choice condition eliminates two difficult requirements of the production condition: recalling the phonological form and inventing a spelling. Because choosing a response is generally easier than producing one (Fischer et al., 1985; Treiman et al., 2015), participants can focus on choosing what they perceive to be the better of the two provided choices instead of spending cognitive

resources recalling the phonological form from the passage. Participants were able to choose the spelling that aligned with their heterography preference in the choice condition.

The production condition was more difficult for participants than the choice condition. Participants struggled to recall the correct phonological form, as evident by the .5 incorrect and no-response rate. Additionally, if the correct phonological form was recalled, it may have been easier for the participant to use a known spelling because inventing a spelling is an inherently difficult task. When a participant remembers the phonological form, not only is the familiar spelling for the homophone activated (Folk, 1999), the participant may consider the multiple ways a critical phoneme (e.g., /i/) can be spelled in English. A participant must consider the established spelling alongside the potential alternatives and settle on one to provide in response to the vocabulary question. If a participant were aware of their heterographic preference throughout the task, they would invent a spelling for each novel homophone, then evaluate it for plausibility. This process of spelling invention and evaluation takes creativity, effort, and time. Instead of thinking of all of the possible ways /i/ can be spelled after hearing /wid/, participants may use the most frequent form that comes to mind and produce ‹weed›. By reusing spellings, participants are perpetuating semantic lexical ambiguity in the language and violating lexical distinctiveness.

Although we addressed many limitations of and successfully replicated Treiman et al. (2015) in Experiment 1, we must acknowledge a major limitation of Experiment 1: the concerning overall rate of incorrect responses in the production condition. Only 50% of the produced responses were able to be used in analyses to compare with choice responses. It is possible that the passages were not informative enough for good incidental vocabulary acquisition. The addition of another sentence containing the novel word might allow participants

in the production condition to remember the phonological form long enough to respond with a plausible answer to the vocabulary question. In a planned follow-up experiment, we intend to implement this change. Participants will hear passages that are three sentences long; the first and last sentence will contain a novel homophone. The additional context and presentation of the novel homophone is expected to improve listeners' ability to recall the novel homophone. We expect to see the same general pattern of results, but with a higher number of plausible spellings in the production condition than in the present Experiment 1. Fewer incorrect responses will allow us to confirm the validity of the present results and corroborate the argument that lexical distinctiveness may not be as important as previously suggested.

# 3. Experiment 2

In Experiment 2, we aimed to extend the findings of Experiment 1 and Treiman et al. (2015) to incidental vocabulary acquisition from silent reading. Silent reading serves as the primary source of vocabulary growth through the lifespan (Long & Shaw, 2000). Between 20 and 60 years of age, the average person learns about 6,000 new words, an average of one every 2 days (Brysbaert et al., 2016). This rate is highest for those who read often, including college students. It is important to understand the processes involved in vocabulary acquisition, including how skilled readers learn the spelling, sounds, and meaning of new words. Of particular interest here is how readers learn homophones. Previous research has focused on how readers learn phonological (Brushnighan et al. 2014; Folk, 1999) and semantic (Maciejewski et al., 2020; Fang et al. 2017; Fang & Perfetti 2018, 2019; Rodd et al. 2012) information about homophones, but less is known about how readers acquire their spellings. The present experiment explores the impact of spelling (i.e., heterographic or homographic) on readers' memory for novel homophones following a silent reading task.

The results of Experiment 1 and Treiman et al. (2015) show that people generally reuse established spellings for novel homophones heard from speech when asked to produce a spelling but select a novel spelling when given a choice. In Experiment 2, we compared people's recall of heterographic and homographic spellings of novel homophones learned from reading to test the value of lexical distinctiveness in vocabulary acquisition from silent reading. Specifically, we examined the proportion of correctly recalled spellings depending on condition (choice or production) and novel homophone spelling (established or novel). Participants read a passage containing a novel homophone and answered a question that elicited the novel homophone's spelling. Participants read half of the novel words with established, homographic spellings (e.g.,

‹noun›) and the other half with novel, heterographic spellings (e.g., ‹nown›). To assess whether lexical distinctiveness is a beneficial feature of writing systems, we tested whether participants would have different rates of correctly recalled spellings for heterographic and homographic novel homophones. We also asked whether participants would recall novel homophone spellings at different rates depending on condition (choice or production). The metalinguistic questionnaire was also included, with the expectation that participants would explicitly express a preference for heterographic spellings, as in Experiment 1 and Treiman et al.

# 3.1 Method

## 3.1.1 Participants

Participants were 55 individuals from the Washington University undergraduate subject pool who received research credit for their participation. All participants had normal or corrected-to-normal vision and hearing. Their mean age was 19.38 years (range 18–22). Two additional individuals participated in the study, but their data were not included because they reported having a language- or reading-related disorder.

## 3.1.2 Stimuli

All participants completed 40 trials, each of which consisted of a passage containing a novel word, a vocabulary question, and a comprehension question. The wording of the passages and questions was identical to that in Experiment 1 (see Table 3.1). We initially programmed the task with a reading rate matching that of the reading rate of the audio recordings in Experiment 1. In a pilot study with 23 participants from the same population, we found that all participants recalled the spelling they saw in each passage with no or very few errors. To avoid ceiling effects, we thus exposed participants to passages and questions for a pre-determined exposure determined by a reading rate of 350 words per minute. This rate is approximately 10% faster

than the upper limit of average adult silent reading rates of fiction text (Brysbaert, 2019), assuming Washington University students will be at the upper range of silent reading rates on average. Passages were presented on the screen for an average of 6.35 seconds (range = 3.94 to 9.09 seconds). Questions were on the screen for an average of 3.2 seconds (range = 2 to 4.5 seconds).

We used the same 80 phonological forms and 160 spellings as Experiment 1 (see Appendix). Because we did not want participants to be exposed to multiple spellings of the same phonological forms (i.e., seeing both ‹noun› and ‹nown› in different passages), we divided the list of phonological forms in half and chose a within-subjects design for the spellings. Each participant saw novel homophones from one of four lists. Half of the items had an established spelling and the other half had a novel spelling. For example, some participants saw 20 established spellings, e.g., ‹noun›, and 20 novel spellings, e.g., ‹cloun›. Other participants instead saw ‹nown› and ‹clown›. There were a total of four counterbalanced lists, ensuring that all items were tested. Approximately equal numbers of participants were randomly assigned to each of the four lists.

Across participants, each passage was seen with two spellings corresponding to one phonological form and two to another, as illustrated in Table 3.1, thus each passage had four versions. We used the same vocabulary and comprehension questions as in Experiment 1. We randomized the order of presentation of the comprehension and vocabulary questions for each passage. For the choice condition vocabulary questions, the participants were offered the established and novel spellings of the novel homophone.

**Table 3.1** Sample of Experiment 2 Passages, Questions, and Alternatives in Choice Condition

| Passage | Questions | | Choice condition alternatives | |
|---|---|---|---|---|
| | Vocabulary | Comprehension | Vocabulary | Comprehension |
| Natalie stared at the nown/noun soaking up the warm sun on the shelf and tried to remember the last time she watered it. She had never noticed how much the spiky plant resembled a tiny round person with a big tuft of orange hair atop his head. | What type of plant does Natalie have? | Where was the plant sitting? | nown<br>noun | on the table<br>on the shelf |
| Maggie reached up to the lowest branch of the clown/cloun and held its tiny yellow berry in her hand. Tears came to her eyes as she reminisced about the days climbing these trees and eating the sweet berries with her sister. | What is the name of the tree that Maggie and her sister used to climb? | What color was the berry Maggie held? | clown<br>cloun | red<br>yellow |

### 3.1.3 Procedure

The general procedure for Experiment 2 was similar to that of Experiment 1. Participants were randomly assigned to one of the four counterbalanced lists and to either the choice or production condition. The experimenter told participants that they would read short passages and answer two questions about each passage. Depending on their assigned condition, participants were instructed to choose or type their response for each question. The experimenter remained in the room while the participant completed two practice trials. Once any questions were answered, the experimenter left the room and closed the door.

Participants sat approximately 60 cm away from a 34.5 cm wide laptop computer screen in a sound-attenuated room. Passages and questions were displayed in white size 24-point, Arial font on a light gray background. After the passage had been presented for its allotted time, a question appeared. Then the participant saw two answer choices or a blank line on which to produce their response, depending on the condition assignment. After answering the two questions for a passage, participants indicated that they were ready to listen to the next passage by pressing a specified key. They were instructed to take breaks as needed and only continue to the next passage when ready.

Participants took 15–20 minutes to complete all 40 passage trials. They then completed several additional tasks. First, they filled out the same optional personal background survey and two-question metalinguistic questionnaire as in Experiment 1. Participants then completed a pronunciation task in which they were audio recorded while reading aloud the list of novel and established spellings from the list they were assigned. Last, participants completed a lexical decision task on the same list of words. The experimenter informed participants that some of the words on the list were invented and some were real English words. Participants circled the words

they believed to be “correctly spelled, real, English words.” The aim of this task was to verify that participants considered the real and the invented spellings as such.

Once all tasks were completed, participants were provided a debriefing document and debriefed verbally. The experimenter summarized the purpose of the study as examining how people learn new words from text. Participants were informed that some of the spellings were invented and some were real English words. They were also informed that the meanings were invented for the purpose of study, and they received a list of the familiar spelled words they read in the experiment. On the debriefing document, each word was accompanied by its proper English definition according to Merriam-Webster (n.d.). For example, in the study ‹noun› was used as a novel word to mean a type of succulent, but on the debriefing sheet we provided the familiar definition of *noun*, “a word … used to identify any of a class of people, places, or things…”. After reviewing the debriefing document, the experimenter answered any questions the participant may have had and then research credit was granted to the participant.

### 3.1.4 Scoring

Responses to comprehension questions were scored as correct or incorrect, as in Experiment 1. A computer scored responses made in the choice condition and two judges scored responses made in the production condition.

Responses to vocabulary questions in the choice condition were scored by a computer program as correct (matched the spelling of the novel homophone read in the passage) or incorrect (did not match the novel homophone spelling in the passage). In the production condition, judges scored the vocabulary question responses as correct or incorrect using the same agreement procedure of Experiment 1. For example, if a participant read ‹noun› as the novel homophone spelling, the correct response to the vocabulary question would be ‹noun›. Incorrect

responses were further sorted into three categories: plausible, illegal, and no-response. Incorrect plausible responses were those that shared a phonological form with the target spelling and indicated correct memory of the phonological form but failure to remember the spelling. For example, ‹nown› or ‹knaun› are incorrect plausible responses for the established spelling ‹noun›. Responses were categorized as incorrect illegal when they did not match the spelling or sound form of the novel homophone spelling presented in the passage. Responses such as ‹rown›, a synonym, or the definition of the novel homophone would be scored as incorrect illegal. Incorrect no-response included blanks and responses such as *???* or *I don't remember*. Incorrect illegal and no-response categories were scored the same as in Experiment 1.

## 3.2 Results

A summary of spelling responses is presented in Table 3.2. In both conditions, whether a participant read an established spelling or novel spelling did not result in better recall. The proportion of correctly chosen established spellings is equal to the proportion of correctly chosen novel spellings. Similarly, the proportion of correctly produced established spellings is equal to the proportion of correctly produced novel spellings. As expected, participants in the choice condition responded correctly more frequently than participants in the production condition.

**Table 3.2** Mean (SD) Proportion of Correct Spellings by Condition and Type of Novel Homophone Spelling in Passage in Experiment 2

| Condition | Novel Homophone Spelling | |
|---|---|---|
| | Established | Novel |
| Choice | .84 (.16) | .85 (.14) |
| Production | .55 (.22) | .56 (.21) |

Data were analyzed with logistic mixed-effects models using the lme4 package in R (Bates et al., 2015). Fixed effects included condition (choice vs. production) and novel homophone spelling (established vs. novel). Model comparisons showed that the interaction of condition and novel homophone spelling made no significant contribution ($\chi^2 = .22$, $p = .64$), so it was not included in the final model. The reference level for condition was choice, and the reference level for novel homophone spelling was established spelling. Random effects terms included intercepts of participant and phonological form. There was a main effect of condition on the log odds of a correct spelling ($\beta = -1.86$, 95% CI [-2.44, -1.29], $z = -6.32$, $p < .001$). There was no significant difference in the log odds of a correct spelling due to the spelling of the novel homophone that was seen in the passage ($\beta = .11$, 95% CI [-.10, .32], $z = 1.06$, $p = .29$). In the choice condition, the predicted probability of correct responses for established spellings was .89 and novel spellings was .90. In the production condition, the predicted probability of correct responses for established spellings was .55 and novel spellings was .58.

We analyzed the types of incorrect responses made by participants in the production condition. When participants saw established spellings for the novel homophone, 38% of responses were illegal and 6% were no-response. Only 0.7% were a novel spelling: ‹weed›/‹wheed›, ‹cane›/‹kane›, and ‹chief›/‹cheif›. When participants saw novel spellings for the novel homophones, 35% of responses were illegal, 4% were no-response, and 4% were established (e.g., saw ‹jaid›, produced ‹jade›). Only two responses (0.5%) were plausible novel spellings that did not match the novel spelling from the passage: ‹wholf›/‹wholfe› and ‹quoat›/‹quoate›.

Participants responded correctly more in the choice condition ($M = .88$, $SD = .07$) than in the production condition ($M = .68$, $SD = .14$), $t(42.75) = 6.73$, $p < .001$. In an exploratory

analysis, we observed a significant correlation of comprehension and vocabulary performance ($r$ = .85, $p$ < .001, Pearson's product-moment) reflecting that participants who correctly recalled comprehension questions also correctly recalled vocabulary questions.

On the metalinguistic questionnaire, 46 participants expressed a preference for novel homophones to be spelled differently from an existing word, 8 participants expressed a preference for the same spelling, and 1 participant expressed no preference. Metalinguistic preference was coded as a factor with three levels (different, same, NA). Adding metalinguistic preference to the model did not add a significant contribution ($\chi^2$ = 1.20, $p$ = .55) suggesting that a preference for heterography does not help predict participants' correct responses. The types of reasons participants provided to justify support for their preference were similar in nature to Experiment 1.

In the pronunciation task, participants correctly pronounced 99.2% of established words and 98.0% of novel words. In the lexical decision task, participants correctly categorized 98.4% of established words as real and 92.7% of novel spellings as invented.

## 3.3 Discussion

If distinctive spellings are beneficial to readers, a novel homophone should be learned better when its spelling is heterographic as opposed to homographic. Lexical distinctiveness, supported by scholars and laypeople alike, would thus be effective in reducing semantic lexical ambiguity and improving clarity for the intended meaning of text. However, in Experiment 2, we found no difference between the rates of recall for heterographic and homographic novel homophones. Within each of the choice and production conditions, established and novel spellings for novel homophones were recalled at equal rates. As expected, most participants expressed an explicit preference for novel spellings (Baker, 1980; Treiman et al., 2015). If

having a heterographic preference impacted behavior, we might have seen those who favor heterography choosing or producing more heterographic spellings. However, we found that participants with an explicit heterographic preference did not choose or produce heterographic spellings more than homographic spellings. Further, if homographic homophones are particularly difficult to learn because of the confusion caused by shared spellings—as participants claimed in justifying their heterographic preference—we would have expected more production responses of novel spellings (e.g., ‹clounn›) when participants saw established spellings (e.g., ‹clown›). Few participants, however, provided novel spellings after reading an established spelling, suggesting that their behavior did not mirror their particular preference for homophones. Participants with heterographic preferences did not have any increased difficulty learning and recalling homographic homophones. Thus, initially thought to be beneficial, the results do not support the idea that lexical distinctiveness is a valuable feature of the English writing system.

To successfully learn new or unfamiliar words, skilled readers must keep track of multiple linguistic features such as the sound, spelling, and meaning of the word. Extensive research focuses on vocabulary acquisition in general, but for homophones specifically, most studies have looked at how readers process phonological information (Brushnighan et al., 2014; Folk 1999; Folk & Morris, 1995; Van Orden, 1987) and access semantic information (e.g., Majewscki et al., 2020; Fang et al., 2017; Fang & Perfetti 2018, 2019; Folk & Morris, 1995; Rodd et al., 2012). Experiment 2 is one of few known studies to focus on skilled readers' learning of novel homophone spellings by directly testing whether readers learn heterographic or homographic novel homophones better. Some studies on novel homophone learning have used tasks that did not collect spontaneous spelling productions (e.g., Folk & Morris, 1995). This is problematic because it does not allow us to make inferences about how people spell new,

unfamiliar homophones. Other studies focusing on learning novel homophone spellings compared novel heterographic homophones (e.g., *skwosh*) to novel nonwords (e.g., *chenth*), but not homographic homophones (e.g., Brusnighan et al., 2014; Folk, 1999). The results of these studies suggest that readers have more difficulty establishing a meaning for novel heterographic homophones than for novel non-homophones. In sum, previous work has shown that readers process and learn the sounds and meanings of new heterographic homophones differently than completely novel words.

The lack of a completely novel word learning comparison is arguably a limitation of the present study. To further test of the value of lexical distinctiveness for readers, we suggest comparing the learning of novel homophones (e.g., ‹noun›/‹nown›) and novel nonhomophones (e.g., ‹joun›/‹jown›) in a task similar to that of Experiment 2. If lexical distinctiveness is valuable to readers, we should see lower recall of critical phoneme spellings (e.g., ‹ow› and ‹ou› for /aʊ/) that result in familiar words than for spellings that result in novel words. For example, ‹ow› in ‹nown› forms a heterographic homophone, while ‹ow› in ‹clown› violates lexical distinctiveness and results in a homographic homophone. For novel nonhomophones with the same critical phoneme, such as ‹jown› and ‹joun›, neither spelling of the critical phoneme forms a familiar word. Thus, if lexical distinctiveness is beneficial, participants should recall spellings avoid homography more than those that result in homography. Specifically, we would expect readers to recall ‹nown› and ‹jown› more than ‹clown›, and ‹cloun› and ‹joun› more than ‹noun›. Assuming the finding that readers are able to learn both spellings of novel homophones equally well replicates, the addition of the novel nonhomophone comparison allows us to clarify whether equal learning of critical phoneme spellings are exclusive to novel homophones. In a pilot study (data collection halted due to Covid-19), we implemented a design that includes

nonhomophones, conceptually replicating Brusnighan et al. (2014). Participants (n = 15) read the same passages and same questions as the present experiment. Half of the words participants saw were homophones (half novel, half established spellings) while the other half of the words participants saw were nonwords. The preliminary results suggest there is no difference in recall of critical phoneme spellings for novel homographic homophones versus novel nonhomophones.

Although the finding that readers chose more spellings correctly than they produced is consistent with previous studies that show recognizing a spelling is easier than recalling it (Fischer, Shankweiler, & Liberman, 1985; Treiman et al., 2015), the high rate of incorrect production responses might lead one to question the validity of our measures. In the production condition, 46% of responses were not plausible spellings of the novel homophones. In Experiment 1, this rate was slightly higher at 50%. One explanation for the high rate of incorrect production responses could be the above-average reading rate we used to present passages to participants. We assumed university students to exceed the average reading speed of 238 words per minute for non-fiction text (Brysbaert, 2019). In future work, one might measure each participant's reading rate and tailor the speed of passage presentation to individual reading rates. This may result in a higher number of correct (or plausible) responses. As suggested for Experiment 1, incorporating additional exposures to the novel homophone may improve encoding and recall for more reliable production condition results. Another approach would be to measure the time participants take to choose or produce a spelling in response to vocabulary questions. While participants were timed in Experiment 2, they were directed to take as much time as needed on each trial. As such, the response times we measured may be inflated and not reflective of actual processing time. Future research should instruct participants to respond as quickly as possible in order to reliably interpret response times. Using a fine-grained measure of

recalling novel homophone spellings from reading (like Folk, 1999) might allow us to see if processing speeds differ in the recall of heterographic versus homographic homophones.

In sum, we found that during silent reading, novel homophones with heterographic spellings are not learned better than those with homographic spellings. The results suggest that novel homographic homophones do not pose a particular problem for skilled readers, despite their own intuitions suggesting so. We found no evidence that lexical distinctiveness offers an advantage for skilled readers learning novel homophones. Some theories describe lexical distinctiveness as a beneficial and foundational principle of a writing system. The present evidence suggests the role of lexical distinctiveness has been overestimated for the English writing system.

# 4.General Discussion

The objective of the current study was to determine the extent to which lexical distinctiveness benefits spellers and readers learning novel homophones. Some scholars suggest that lexical distinctiveness is a desirable feature of writing systems as it improves clarity by reducing semantic ambiguity (Rutkowska & Rossler, 2012; Venezky, 1999). By consistently offering different spellings when words sound the same, a writing system can avoid homographic homophones that would cause confusion. We asked whether spellers would choose and produce heterographic spellings more than homographic spellings for novel homophones presented auditorily and whether readers would learn heterographic homophones better than homographic homophones.

The results of our two studies suggest that, although spellers and readers see a benefit of heterography when thinking about it, heterography does not offer an advantage in learning novel homophones. When English language users were asked, they overwhelmingly stated that heterographic spellings are better for novel homophones when they are introduced into a language. These responses were found in the current experiments, Treiman et al. (2015), and Baker (1980). Despite seeing the benefit, people did not behave accordingly. Spellers produced more homographic spellings than heterographic spellings. While participants did tend to choose heterographic spellings when given a choice, the results of the production task better reflect what spellers do in the real world when faced with writing down a novel word that they have just heard. In addition, readers in both the choice and production condition did not learn heterographic spellings any better than homographic spellings. Our findings suggest that lexical distinctiveness may not be as valuable to spellers and readers as scholars and laypeople believe.

There are several potential explanations for the explicit endorsement of lexical distinctiveness but lack of an advantage to heterography in learning novel homophones.

One possibility is that the English writing system does not distinguish homophonous words systematically. Some linguists, in fact, have suggested that the majority of homophones are not heterographic and question the extent to which distinctiveness should be considered a fundamental principle of English (Carney, 1994; Ryan, 2016; Ziff, 1967). In response to these ideas, Berg and Aronoff (2020) sought to determine the extent to which the English writing system distinguishes homophones orthographically. They used the CELEX corpus, a data set that contains the orthography, phonology, morphology, syntax, and frequency for words in English (Baayen et al. 1995) and printed dictionaries, to gather morphologically simple homophones. Their list of 608 homophonous forms is the most comprehensive record of heterographic and homographic homophones to date. If heterography truly is important and valuable to a writing system, as suggested by the principle of lexical distinctiveness and language users (Baker, 1980; Treiman et al., 2015), one would expect all or most homophones to have heterographic spellings. However, 281 of the 608 forms, 46%, had the same spelling. The patterns used to differentiate heterographic homophones are not used consistently. If heterography were a major principle of English spelling, then we would expect to see patterns of differentiation used to a much higher degree. For example, one pattern used to differentiate homophones with /i/ is to spell the vowel as ‹ee› or ‹ea›, as in *leak/leek, heel/heal,* and *deer/dear*. However, the writing system does not use this pattern to differentiate the word *treat*—meaning to deal with someone in a certain way—from its homophonous counterpart, *treat*—meaning an object given as a reward. The research does not support a principle of lexical distinctiveness by suggesting that the English writing system does not systematically use distinct spellings to distinguish homophonous words.

In English, using heterography to distinguish homophones is not systematic and can be primarily explained by sound change over time (Berg & Aronoff, 2020; Carney 1994). Present-day homophones were not always homophones thanks to historical accidents. That is, sounds change but spellings do not. Two words with distinct spellings and distinct sounds undergo a merger such that the sounds become indistinguishable. For example, prior to the 18th century, ‹leak› and ‹leek› had distinct phonological forms in English. At some point in the 18th century, the sounds made by the letters ‹ea› and ‹ee› merged into what is presently /i/, resulting in heterographic homophones. Orthographic distinction of some homophones and not others is primarily due to historical accident, not systematic differentiation of words that sound the same and have different meanings.

Because heterography is accidental and not a regular pattern in English, it seems odd that language users and theories endorse it. Endorsing heterography does not align with most of the behaviors observed in the present experiments; spellers produced few heterographic homophones, and readers correctly recalled the spellings of both heterographic and homographic homophones. One possible explanation lies in the discrepancy between what an individual knows versus having the ability to express what they know. Linguistic knowledge is theorized to be acquired through a process called statistical learning, the tracking of statistical regularities in language and other domains (see Romberg & Saffran, 2010, for a review). People are implicitly aware of the patterns governing their speaking, reading, and writing. However, people do not always have the metalinguistic awareness to explain the patterns they regularly use to communicate effectively. Treiman and Wolter (2018) examined whether participants were influenced by certain patterns when deciding whether to double the medial consonant of words (e.g., *spinet* vs *spinnet*). They found that participants were influenced by particular features of

the word, yet only 23% of participants could explicitly state the rule that governed their consonant-doubling decisions. The researchers suggest that knowledge acquisition through statistical learning is often not accessible to conscious awareness. The results of the present experiments are not surprising, if it is the case that, through statistical learning, language users have learned that heterography is not a systematic feature of the English writing system. However, this explanation does not clarify why people chose heterographic homophones in the spelling experiment. In addition to revising existing theories supporting lexical distinctiveness as a principle in English, future work should continue to address the discrepancy between people's explicit belief and the linguistic reality that drives their behaviors.

A second possibility that may explain the results of the current study lies in the simple ease of reusing spellings. By reusing easily accessible linguistic units (in this case, familiar spellings), people can follow the principle of least effort: that humans naturally take the path of least resistance in many domains (Zipf, 1949). In the spelling choice condition, we gave participants the path requiring least effort by providing them with two plausible choices. In production condition, however, participants had to find the path of least resistance, which was to produce an established spelling. Producing established spellings violated lexical distinctiveness and deviated from participants' explicit endorsement of heterography. In the reading task, participants readily learned homographic spellings. The behaviors in the spelling production condition and reading experiment introduced or maintained semantic ambiguity which is often thought of as a detriment to a language. Some linguists, however, suggest semantic ambiguity may be a positive feature of natural language that has evolved to encourage efficient communication (Piantadosi et al., 2012). By reusing spellings, participants were able to contribute to the task efficiently. The current results support the idea that semantic ambiguity is

beneficial and offer further evidence that the importance of lexical distinctiveness has been overestimated in explaining what makes a writing system ideal.

A final possibility that may work against lexical distinctiveness is the muddy divide between polysemes and homophones. Additional meanings to familiar words are continuously added to our language. *Ghost* coming from a Gen Z mouth does probably not refer to a spooky spirit but to someone abandoning a relationship. If you hear someone say, "This party is *lit*," chances are that they are not talking about the illumination in the room but how much fun they are having. Language users often update their lexicon upon hearing or seeing new words. The seemingly new meanings of *ghost* and *lit* are not completely unrelated to the original meanings, suggesting these hip slang words are polysemes, or semantically ambiguous words with one orthographic representation and multiple senses, and not homophones. This realization leads to a bigger question: where is the dividing line between homographic homophones and polysemes?

Perhaps the prevalence of polysemy in part explains people's acceptance and use of homographic homophones despite their explicit endorsement of heterography. At least half—some linguists say "most"—of all English words are semantically ambiguous, that is, they are either polysemes or homographic homophones (Eddington & Tokowicz, 2015; Ziff, 1967). Most are polysemes (Copestake & Briscoe, 1995; Jackendoff, 2002; Murphy, 2002; Pustejovsky, 1995). Further, polysemes and homographic homophones are represented and processed differently (Eddington & Tokowicz, 2015; Frazier & Rayner, 1990; Frisson & Pickering, 1999; Pickering & Frisson, 2001; Klepousniotou et al., 2008). In order to explain the difference in processing, Klepousniotou et al. (2008) suggested that semantic ambiguity be defined along a continuum of meaning overlap. On one end of the continuum is metonymy or polysemes with highly overlapping meanings, such as *chicken* referring to the animal or the meat of the animal.

On the other end of this ambiguity spectrum lie homographic homophones, where the two meanings of one orthographic form have little or no overlap, such as *squash* the racket sport and the vegetable. In the middle of the spectrum are polysemes with moderately overlapping meanings. These tend to be metaphorical, such as a *book* being a physical, bound collection of pages or the imaginative content that an author types into a word processor.

Where a word is on the continuum influences a reader's comprehension of an ambiguous word (Klepousniotou et al., 2008). To determine the effects of context and meaning overlap on the processing and memory of ambiguous words, Klepousiotou et al. (2008) had participants judge whether ambiguous words made sense. Participants read phrases containing a modifier, which served to bias the meaning or provide neutral context, and an ambiguous word, such as *tasty chicken, baby chicken,* or **** *chicken.* Overall, the results indicate that high-overlap words are processed more quickly and accurately than moderate- and low-overlap words. High-overlap words are processed most easily because, the authors suggest, polysemes with high semantic overlap have a single, core representation for the two words. The core representation is active irrespective of context. For example, when one reads *chicken* as part of a menu item, the meaning for clucking *chicken* is also activated. Thus, for ambiguous words with high semantic overlap, processing is not impaired as it is with homographic homophones. Homographic homophones, having very little or no semantic overlap have distinct meaning representations in the lexicon, result in processing delays or difficulties when they are encountered in text.

Because homographic homophones and polysemes seem to be processed differently (Klepousniotou et al., 2008), it is important to consider whether our homographic stimuli were perceived as polysemes instead of homophones. Although we attempted to make the semantic relationship as small as possible, perhaps our participants took the new words as somehow

related to or extensions of the original meaning. Participants may have benefitted by interpreting homophones as polysemes. For example, in one passage, a boy piles his plate high with slices of his favorite fruit, ‹brain›/‹brane›, which is being served on a buffet. Upon hearing or reading /bren/, the familiar meaning for *brain* as a neurological organ is activated (Klepousniotou et al., 2008; Folk, 1999; Folk & Morris, 1995; Brusnighan et al., 2014). Although it is not immediately obvious that there is any relationship between the concept we provide for /bren/ and the organ *brain*, participants may search for relationships in order to make sense of a novel meaning for the seemingly familiar word. The brain in a body is similar to fruit in being edible, fleshy, and sliceable. If we consider *brain* to be polysemous with moderate or high semantic overlap, its processing would be different than that of a truly homophonous word. One way to identify where our stimuli fall along the continuum of meaning overlap (Klepousniotou et al., 2008) would be to quantify the semantic distance between the novel homophone meaning and its real word counterpart. If this could be tested, it would allow us to examine the extent to which semantic overlap affects participants' spelling and recall of novel homophones. However, polysemy fails to explain the low rates of homographic spellings in the choice task of Experiment 1. That is, if polysemy were the entire explanation, participants in Experiment 1 would have chosen ‹brain› instead of ‹brane› to spell the novel homophone.

# 4.1 Conclusions

The purpose of the current study was to examine the value of lexical distinctiveness for spellers and readers. When given a choice of spelling for a novel homophone, people reliably choose heterographic spellings (Experiment 1; Treiman et al., 2015). However, people do not spontaneously produce heterographic spellings for novel homophones, nor do they learn heterographic homophones better than homographic homophones in reading (Experiment 2). We

discussed several factors that may work against lexical distinctiveness: the distribution of heterographic homophones in English, the simple ease of reusing spellings, and the muddled division between polysemes and homophones. Taken together, the results of these two studies suggest that lexical distinctiveness is not and should not be endorsed as being advantageous to the processing of or memory for homophones.

# References


Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database (CD-ROM). Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.

Baker, R. G. (1980). Orthographic awareness. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 51–68). Academic Press.

Bates D, Mächler M, Bolker B, Walker S (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48. http://doi.org/gcrnkw

Berg, K., & Aronoff, M. (2020). Unique lexeme spellings. *Morphology*, 2–18.

Brusnighan, S. M., Morris, R. K., Folk, J. R., & Lowell, R. (2014). The role of phonology in incidental vocabulary acquisition during silent reading. *Journal of Cognitive Psychology, 26*(8), 871–892. http://doi.org/dr48

Brysbaert, M. (2019). How many words do we read per minute? A review and meta-analysis of reading rate. *Journal of Memory and Language, 109*, Article 104047. http://doi.org/gf93nd

Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. (2016). How many words do we know? Practical estimates of vocabulary size dependent on word definition, the degree of language input and the participant's age. *Frontiers in Psychology*, *7*(JUL), 1–11. http://doi.org/d4hd

Carney, E. (1994). *A survey of English spelling*. Routledge.

Copestake, A., & Briscoe, T. (1995). Semi-productive polysemy and sense extension. *Journal of semantics*, *12*(1), 15–67. http://doi.org/cg6vhc

Cummings, D. W. (1988). *American English spelling: An informal description.* Johns Hopkins University Press.

Eddington, C. M., & Tokowicz, N. (2014). How meaning similarity influences ambiguous word processing: the current state of the literature. *Psychonomic Bulletin and Review*, *22*(1), 13–37. http://doi.org/f62msx

Fang, X., & Perfetti, C. A. (2018). Learning new meanings for known words: Perturbation of original meanings and retention of new meanings. *Memory and Cognition, 47*(1), 130–144. http://doi.org/gft5hw

Fang, X., Perfetti, C., & Stafura, J. (2017). Learning new meanings for known words: Biphasic effects of prior knowledge. *Language, Cognition and Neuroscience, 32*(5), 637–649. http://doi.org/dr46

Fischer, F. W., Shankweiler, D., & Liberman, I. Y. (1985). Spelling proficiency and sensitivity to word structure. *Journal of Memory and Language, 24*(4), 423–441. http://doi.org/cm7vfb

Folk, J. R. (1999). Phonological codes are used to access the lexicon during silent reading. *Journal of Experimental Psychology: Learning Memory and Cognition, 25*(4), 892–906. http://doi.org/dr5vqh

Folk, J., & Morris, R. K. (1995). Multiple lexical dodes in reading: Evidence from eye movements, naming time, and oral reading. *Journal of Experimental Psychology: Learning Memory and Cognition, 21*(6), 1412–1429. http://doi.org/bthtks

Frazier, L., & Rayner, K. (1990). Taking on semantic commitments: Processing multiple meanings vs. multiple sense. *Journal of Memory and Language, 29*(2), 181–200. http://doi.org/chb6kn

Frisson, S., & Pickering, M. J. (1999). The processing of metonymy: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(6), 1366–1383. http://doi.org/csq6ch

Jackendoff, R. (2002). *Foundations of language.* Oxford University Press.

Kessler, B., & Treiman, R. (2015). Writing systems: Their properties and implications for reading. In A. Pollastek & R. Treiman (Eds.), *The Oxford Handbook of Reading*. (pp. 10–25). http://doi.org/dr45

Klepousniotou, E., Titone, D., & Romero, C. (2008). Making sense of word senses: The comprehension of polysemy depends on sense overlap. *Journal of Experimental Psychology: Learning Memory and Cognition, 34*(6), 1534–1543. http://doi.org/bxn3qd

Maciejewski, G., Rodd, J. M., Mon-Williams, M., & Klepousniotou, E. (2020). The cost of learning new meanings for familiar words. *Language, Cognition and Neuroscience, 35*(2), 188–210. http://doi.org/d4hc

Merriam-Webster. (n.d.) *Merriam-Webster.com dictionary*. Retrieved Summer 2019, from https://www.merriam-webster.com/

Murphy, G. (2002). *The big book of concepts.* MIT Press.

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 51*(1), 195–203. http://doi.org/gft89w

Piantadosi, S. T., Tily, H., & Gibson, E. (2012). The communicative function of ambiguity in language. *Cognition, 122*(3), 280–291. http://doi.org/fx4mkj

Pickering, M. J., & Frisson, S. (2001). Processing ambiguous verbs: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(2), 556–573. http://doi.org/cnmf2p

Pustejovsky, J. (1995). *The generative lexicon*. MIT Press.

Rodd, J. M., Berriman, R., Landau, M., Lee, T., Ho, C., Gaskell, M. G., & Davis, M. H. (2012). Learning new meanings for old words: Effects of semantic relatedness. *Memory and Cognition*, *40*(7), 1095–1108. http://doi.org/d4hb

Rodd, J., Gaskell, G., & Marslen-Wilson, W. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language, 46*(2), 245–266. http://doi.org/b8v8xf

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(6), 906–914.

Rutkowska, H., & Rössler, P. (2012). Orthographic variables. In J. M. Hernández-Campoy & J. C. Conde-Silvestre (Eds.), *The Handbook of Historical Sociolinguistics* (pp. 211–236). Blackwell. http://doi.org/dr44

Ryan, D. (2016). Linguists' descriptions of the English writing system. In V. Cook & D. Ryan (Eds.), *Routledge handbook of the English writing system* (pp. 41–64). Routledge.

Small, S., Cottrell, G., & Tanenhaus, M. (1988). Preface. In S. Small, G. Cottrell., & M. Tanenhaus (Eds.), *Lexical Ambiguity Resolution: Perspectives from Psycholinguistics, Neuropsychology, and Artificial Intelligence* (1st ed., pp. 3–5). Morgan Kaufmann Publishers.

Treiman, R., Seidenberg, M. S., & Kessler, B. (2015). Influences on spelling: Evidence from homophones. *Language, Cognition and Neuroscience*, *30*(5), 544–554. http://doi.org/dr5c

Treiman, R., Wolter, S. (2018). Phonological and graphotactic influences on spellers' decisions about consonant doubling. *Memory & Cognition, 46,* 614–624. http://doi.org/d6pd

Van Orden, G. C., Johnston, J. C., & Hale, B. L. (1988). Word identification in reading proceeds from spelling to sound to meaning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(3), 371–386. http://doi.org/frw8tc

Venezky, R. L. (1999). *The American way of spelling: the structure and origins of American English orthography*. Guilford Press.

Zipf, G. (1949). *Human behavior and the principle of least effort*. Addison-Wesley.

Ziff, P. (1967). Some comments on Mr. Harman's confabulations. *Foundations of Language*, *3*(4), 403–408.

# Appendix

*The homophone phonological form is presented first, followed by the established and novel spellings in parentheses. Sets of items on the same line share a critical phoneme.*

/taɪt/ (tight, tite)
/faɪt/ (fight, fite)
/klaʊn/ (clown, cloun)
/red/ (raid, rade)
/ʃel/ (shale, shail)
/tret/ (trait, trate)
/ler/ (lair, lare)
/kren/ (crane, crain)
/bɜrd/ (bird, berd)
/fedʒ/ (phage, fage)
/bik/ (beak, beek)
/wid/ (weed, wead)
/rif/ (reef, reaf)
/sit/ (seat, seet)
/fir/ (fear, feer)
/dʒoʊk/ (joke, joak)
/loʊd/ (load, lode)
/roʊp/ (rope, roap)
/ræθ/ (wrath, rath)
/wʊlf/ (wolf, wholf)
/hwaɪt/ (white, wite)
/kaɪt/ (kite, kight)
/kraʊd/ (crowd, croud)
/dʒel/ (jail, jale)
/der/ (dare, dair)
/det/ (date, dait)
/frem/ (frame, fraim)
/ken/ (cane, cain)
/fɜrm/ (firm, ferm)
/foʊn/ (phone, fone)
/fil/ (feel, feal)
/lif/ (leaf, leef)
/sid/ (seed, sead)
/piʧ/ (peach, peech)
/flit/ (fleet, fleat)
/moʊd/ (mode, moad)
/goʊl/ (goal, gole)
/floʊt/ (float, flote)
/rɪf/ (riff, wriff)
/wit/ (wheat, weat)

/spaɪt/ (spite, spight)
/baɪt/ (bite, bight)
/naʊn/ (noun, nown)
/dʒed/ (jade, jaid)
/nel/ (nail, nale)
/kret/ (crate, crait)
/ker/ (care, cair)
/sten/ (stain, stane)
/nɜrd/ (nerd, nird)
/fek/ (fake, phake)
/gik/ (geek, geak)
/mid/ (mead, meed)
/ʧif/ (chief, cheef)
/twit/ (tweet, tweat)
/ʧir/ (cheer, chear)
/soʊk/ (soak, soke)
/koʊd/ (code, coad)
/soʊp/ (soap, sope)
/ræg/ (rag, wrag)
/wʊp/ (whoop, woop)
/fraɪt/ (fright, frite)
/flaɪt/ (flight, flite)
/laʊd/ (loud, lowd)
/kel/ (kale, kail)
/ʧer/ (chair, chare)
/bet/ (bait, bate)
/klem/ (claim, clame)
/bren/ (brain, brane)
/tɜrm/ (term, tirm)
/fɔrk/ (fork, phork)
/sil/ (seal, seel)
/bif/ (beef, beaf)
/bid/ (bead, beed)
/spiʧ/ (speech, speach)
/trit/ (treat, treet)
/toʊd/ (toad, tode)
/moʊl/ (mole, moal)
/kwoʊt/ (quote, quoat)
/rɪst/ (wrist, rist)
/wiv/ (weave, wheave)