Neutralization and Homophony Avoidance in Phonological Learning

Sora Heng Yin

New York University

James White

University College London

Abstract

Previous research has suggested that homophony avoidance plays a role in constraining language change; in particular, phonological contrasts are less likely to be neutralized if doing so would greatly increase the amount of homophony in the language. Most of the research on homophony avoidance has focused on the history of real languages, comparing attested and unattested (hypothetical) phonological changes. In this study, we take a novel approach by focusing on the language learner. Using an artificial language learning paradigm, we show that learners are less likely to acquire neutralizing phonological rules compared to non-neutralizing rules, but only if the neutralizing rules create homophony between lexical items encountered during learning. The results indicate that learners are biased against phonological patterns that create homophony, which could have an influence on language change. The results also suggest that lexical learning and phonological learning are highly integrated.

Keywords: phonology, language change, learning bias, functional load, neutralization, artificial language

1. Introduction

A crucial function of speech sounds is that they allow speakers to contrast words. For instance, the English words *tab* and *dab* differ only in their initial sounds ([t] vs. [d]), suggesting that /t/ and /d/ represent two basic sound categories, or phonemes, in English (Hayes, 2009). However, the phonological rules of a language sometimes result in the neutralization of phonemic contrasts. For example, the flapping rule in American English affects both /t/ and /d/, changing them into an alveolar flap [ɾ] between vowels when the second vowel is unstressed; as a result, the distinct lexical items *pat* [pæ**t**] and *pad* [pæ**d**] have the same phonetic realization when the suffix *–ing* is added (i.e., *patting* and *padding* are both pronounced as [pæ**ɾ**ɪŋ]).[1]

We can distinguish two levels of neutralization. First, there can be neutralization at the lexical level, such as when flapping results in the same pronunciation of the words *patting* and *padding.* This level of neutralization could be called derived homophony (Silverman, 2012, p. 4). Second, there is neutralization within the phonological system. Neutralizing phonological rules, by definition, eliminate a contrast between two (or more) phoneme categories; for instance, the contrast between /t/ and /d/ is lost in American English in flapping contexts because both are realized as the same sound, [ɾ]. These two levels of neutralization (i.e., phonological and lexical) are clearly related, but are partially distinct. The application of a neutralizing phonological rule may result in homophony, or it may not. For

---

[1] Strictly speaking, flapping in American English is not completely neutralizing from a phonetic perspective. Subtle acoustic differences remain between /t/ and /d/ after flapping, including a small (but statistically reliable) difference in the duration of the preceding vowel. However, adult speakers are unable to reliably distinguish flapped /t/ and /d/ in perception (Herd et al., 2010).

instance, flapping occurs in the word *getting* [ɡɛɾɪŋ], but there is no lexical neutralization in this particular case because *ged* and *gedding* are not existing words of English. In sum, while neutralizing phonological rules have the *potential* to create homophony, the amount of *actual* homophony that they create can vary depending on the contents of the lexicon.[2]

Neutralization poses a challenge for our understanding of language change and typology. On one hand, neutralization creates ambiguity, which reduces the communicative efficiency of a language. On the other hand, neutralization is not uncommon in the world's languages (for an overview, see Silverman, 2012). Are there pressures against the development of neutralizing rules given that they increase ambiguity? If so, which mechanisms are responsible for such pressures, and which factors influence whether a neutralizing rule will eventually develop in a language?

According to the FUNCTIONAL LOAD HYPOTHESIS (Martinet, 1952; Hockett, 1967; King, 1967; Wedel, Kaplan, & Jackson, 2013; see also earlier work: Gilliéron, 1918; Jakobson, 1931; Mathesius, 1931), the likelihood of two phonemes being neutralized over the course of language change depends on the amount of information that they carry: pairs of phonemes that play a greater role in contrasting meaning in a language (i.e., have a higher functional load) are less likely to be neutralized over time. Applied to phonological rules, the theory predicts that a neutralizing rule resulting in a large amount of homophony is less likely to be adopted into a language than one resulting in little homophony, all else being equal.

---

[2] We use the term 'rule' throughout the paper as a straightforward way of referring to a context-sensitive change from one sound to another that could be generalized to novel cases. We are not suggesting that the cognitive structure of the grammar is organized as in traditional rule-based theories of phonology (e.g., Chomsky & Halle, 1968).

A prediction that emerges from the functional load hypothesis is that the amount of homophony created amongst *existing* lexical items should affect the likelihood of a neutralizing rule being adopted or applied. As Kaplan (2011) points out, this outcome is not necessarily expected under all models of phonology. While models of phonology with connections to exemplar theory have embraced the idea that the lexicon and the phonological system are highly integrated (Hay, 2000; Pierrehumbert, 2001, 2002; Bybee, 2002; Ernestus & Baayen, 2003; Wedel, 2004, 2006, 2007; Ussishkin & Wedel, 2009), traditional generative models of phonology have typically maintained a modular view of the lexicon and the phonology. In generative models, the domain of phonological analysis is considered to be the set of possible words in a language, not the set of existing words in the lexicon (e.g., see Halle, 1962; Padgett, 2003, 2009; Ní Chiosáin & Padgett, 2009).

While the functional load hypothesis is intuitively appealing, testing whether homophony avoidance actually affects the likelihood of neutralization has proven difficult. Many factors play a role during language change. Though we often know which phonological changes have occurred in a language, it is difficult to identify which other hypothetical changes could have occurred and determine why they did not occur (e.g., see discussion in Kaplan, 2011). In the current study, we take a novel approach to this issue by focusing directly on the language learner. On the basis of two artificial language experiments, we show that learners are less likely to acquire neutralizing rules than non-neutralizing rules, but only when the neutralizing rules create homophony. The results provide support for the functional load hypothesis and suggest that a learning bias is one mechanism through which homophony avoidance could shape phonological systems.

In the following sections, we review previous literature on homophony avoidance, provide a brief overview of the artificial language learning paradigm, and then introduce the experiments of the current study.

1.1. Homophony avoidance in language change

One approach to testing whether homophony avoidance plays a role in language change is to compare the amount of homophony created by attested phonological changes and hypothetical changes that did not occur. Relevant phonological changes include complete phoneme mergers (neutralization in all contexts) and the adoption of new phonological rules which neutralize contrasts in certain contexts. Several studies taking this approach have concluded that homophony avoidance plays a significant role in constraining phonological change. For instance, Silverman (2010) argued that even though Korean has several neutralizing rules, these rules create less homophony than other comparable (but non-occurring) rules would. Expanding on this research, Kaplan (2011) used Monte Carlo simulations to show that the attested neutralizing rules in Korean produce fewer homophones than we would expect based on the distribution of homophony created by randomly generated sets of comparable unattested rules.[3] Wedel et al. (2013) conducted a statistical analysis of neutralizing diachronic changes (mergers and new neutralizing rules) in nine languages. They found that the number of minimal pairs distinguished by a pair of phonemes was a significant

---

[3] This was true for nouns. The results for verbs were more equivocal.

predictor of whether the pair would be neutralized, even taking into account other factors such as overall phoneme probability.

Other researchers, however, have raised doubts about the claim that functional load, or homophony avoidance, plays a significant role in phonological change. King (1967) analyzed the history of phoneme mergers in several Germanic languages and concluded that functional load is, at best, one of the least important factors in sound change. Sampson (2013) likewise argued that the history of phonological changes from Middle Chinese to Modern Chinese shows an increase in the amount of homophony between morphemes, which is inconsistent with homophony avoidance being a major factor in language change (see Kaplan, 2015 for a reply). Thus, the role of homophony avoidance in sound change remains an unsettled issue.

These disagreements underscore some major challenges facing this line of research. It is difficult to determine precisely how much homophony should be considered inconsistent with the functional load hypothesis. Even if we interpret the theory as stating that attested changes should produce less homophony than is expected by chance, it is difficult to identify which set of hypothetical changes should be used for calculating the ‘expected’ amount of homophony. Many interacting factors affect the likelihood that a phonological change will occur. Moreover, some phonological changes are inherently more likely to occur than others, independent from any effect of homophony avoidance. Defining a set of comparable hypothetical rules is far from a trivial endeavor (e.g., see discussion in Kaplan, 2011).

To show definitively that a hypothetical change failed to occur in a language due to homophony avoidance, one would need to demonstrate that the hypothetical change would have occurred if all conditions were identical except that it created less homophony (Kaplan, 2011). Blevins & Wedel (2009) cover cases in the history of natural languages where this was hypothesized to have occurred, but confirming that homophony avoidance was the cause in any particular case is challenging, particularly given the potential influence of other factors such as analogical change. In the history of natural languages, it is difficult to determine with confidence what speakers would have done if circumstances had been different.

However, using a tightly controlled artificial language learning experiment, we can indeed probe what learners will do under minimally different circumstances. Specifically, we can provide learners with the same set of rules to learn, varying only the amount of homophony created by those rules. This is the approach we took in the current study, allowing us to directly test whether homophony avoidance makes learners less likely to adopt neutralizing phonological rules.

1.2. Exploring learning biases using artificial language learning

The language learner plays an important role in shaping language change. Over many generations, even subtle learning biases can shift languages in certain directions (Kirby, Smith, & Brighton, 2004; Kalish, Griffiths, & Lewandowsky, 2007; Reali & Griffiths, 2009; Culbertson, Smolensky, & Legendre, 2012; White, 2017). Understanding the biases that learners bring to the language learning process is therefore a key component of explaining language change and typology. Artificial language learning experiments have emerged as a useful framework for probing learning biases with adults, children, and infants (Onishi, Chambers, & Fisher, 2002; Chambers, Onishi, & Fisher, 2003; Saffran & Thiessen, 2003; Newport & Aslin, 2004; Seidl & Buckley, 2005; Peperkamp, Skoruppa, & Dupoux, 2006; Wilson, 2006; Cristià & Seidl, 2008; Moreton, 2008; Finley & Badecker, 2009; Carpenter, 2010; Smith & Wonnacott, 2010; Skoruppa & Peperkamp, 2011; Finley, 2011; Wonnacott, 2011; Baer-Henney & van de Vijver, 2012; Culbertson et al., 2012; Finley & Badecker, 2012; White, 2014; White & Sundara, 2014; Culbertson & Newport, 2015; Fehér, Wonnacott, & Smith, 2016; Finley, 2017; Wonnacott, Brown, & Nation, 2017; for reviews see Gomez & Gerken, 2000; Moreton & Pater, 2012a, 2012b). These studies have demonstrated that participants are able to learn novel linguistic patterns (including phonological rules) after brief exposure to an artificial language in the lab.

Unlike in natural language learning, artificial language experiments allow the researcher to have maximal control over the amount and type of input received by learners. Participants receive input from one of several minimally different pseudo-languages, or input that is ambiguous between multiple analyses. If participants' responses at test are biased in a way that cannot be explained by their input—for instance, if they learn one pattern better than another pattern given equal input for both patterns—then the results provide evidence for a learning bias.

1.3. Overview of current study

We designed two artificial language experiments to test whether neutralization and homophony avoidance affect how well participants learn novel phonological rules.

Experiment 1 tested the hypothesis that learners are biased against neutralizing rules compared to non-neutralizing rules. We exposed learners to one of two artificial languages exhibiting the same four phonological rules, two of which were neutralizing and two of which were non-neutralizing (counterbalanced between the groups). The neutralizing rules in Experiment 1 were both phonologically neutralizing and lexically neutralizing; specifically, the neutralizing rules created homophony (between lexical items encountered during exposure) in 50% of the cases where they applied. To preview the results, we found that the neutralizing rules were indeed learned more poorly than the non-neutralizing rules. In Experiment 2, we explored the extent to which this bias was driven by homophony avoidance by manipulating the amount of homophony created by the neutralizing rules during learning. To maximize the distinction, we used extreme values for homophony in Experiment 2: neutralizing rules resulted in homophony in either 0% or 100% of the cases where the rules applied.

Together, the two experiments allow us to examine the effects of both phonological neutralization and lexical neutralization on phonological learning. In all cases, neutralizing rules resulted in phonological neutralization whereas non-neutralizing rules did not. However, the amount of lexical neutralization, or homophony, varied across the groups that we tested (50% in Exp. 1; 0% or 100% in Exp. 2). If homophony avoidance affects the likelihood that learners will acquire a phonological rule, then we should see a stronger bias against neutralizing rules when they create homophony between observed lexical items. On the other hand, if phonological learning is not sensitive to the actual contents of the lexicon, then the (phonologically) neutralizing rules should be harder to learn regardless of how much homophony they create.

2. Experiment 1

Experiment 1 was designed to test whether learners have a bias against neutralization when presented with equal evidence for both neutralizing and non-neutralizing rules in their input. Participants were randomly assigned to one of two exposure groups, Language A or Language B. Table 1 summarizes the sound inventories and phonological rules of the two languages.

Table 1. Summary of Language A and Language B.

	Language A	Language B
Critical alternating phonemes	/t, d, s, z/	/t, d, s, z/
Critical non-alternating phonemes	/tʃ, dʒ/	/ʃ, ʒ/
Neutralizing rules	/t, d/ → [tʃ, dʒ] / __i	/s, z/ → [ʃ, ʒ] / __i
Non-neutralizing rules	/s, z/ → [ʃ, ʒ] / __i	/t, d/ → [tʃ, dʒ] / __i
Filler phonemes (non-alternating)	/p, b, k, g, f, v/	/p, b, k, g, f, v/

In the exposure phase, participants were presented with novel singular/plural word pairs (e.g., singular [zudap], plural [zudapi]), where the plural was marked by an *–i* suffix attached to the singular form. Crucially, some of the word pairs exhibited phonological alternations involving palatalization of the final stem consonant (e.g., singular [tusut], plural [tusutʃi]).[4] Participants in both groups were exposed to the same four rules, /t, d, s, z/ → [tʃ, dʒ, ʃ, ʒ]

---

[4] Transcriptions are in the International Phonetic Alphabet (IPA). To be explicit: [tʃ] = 'tch' in *ma<u>tch</u>*; [ʃ] = 'sh' in *ma<u>sh</u>*; [dʒ] = 'dge' in *ba<u>dge</u>*; [ʒ] = 's' in *vi<u>s</u>ion*; [i] = 'ee' in *f<u>ee</u>*; [u] = 'oo' in *z<u>oo</u>*. [a] was produced similar to Spanish 'a'. Other symbols were pronounced as in English.

before /i/. In Language A, the sounds /tʃ, dʒ/ also had phonemic status, making the [t, d] → [tʃ, dʒ] rules neutralizing; the /s, z/ → [ʃ, ʒ] rules were non-neutralizing because [ʃ, ʒ] occurred only as allophonic variants of the phonemes /s, z/. In Language B, /ʃ, ʒ/ were phonemes rather than /tʃ, dʒ/, making the /s, z/ → [ʃ, ʒ] rules neutralizing and the /t, d/ → [tʃ, dʒ] rules non-neutralizing. Thus, all participants learned the same four rules, /t, d, s, z/ → [tʃ, dʒ, ʃ, ʒ], but the two rules that were neutralizing in Language A were non-neutralizing in Language B, and *vice versa*. This counterbalancing measure ensured that any differences observed in learning must be due to whether the rules were neutralizing or non-neutralizing, not due to an inherent property of the rules themselves. In addition, the two languages shared an identical set of filler phonemes /p, b, k, g, f, v/, which did not change when followed by /i/.

In the test phase, participants completed a forced-choice task on a set of trained items and a set of novel items. After hearing the singular form (e.g., [dazat]), participants were presented with two plural options, a palatalized option ([dazatʃi]) and a non-palatalized option ([dazati]). They had to choose the correct plural option by pressing a button.

To summarize, the design of Experiment 1 was a 2 x 2 within-subjects design. The independent variables were Item Type (Trained or Novel) and Rule Type (Neutralizing or Non-neutralizing). Exposure Language (Language A or Language B) was an additional counterbalancing variable (between-subjects). The dependent variable was accuracy in terms of selecting the appropriate plural item at test.

If learners are biased against learning neutralizing rules, participants should have higher accuracy when applying non-neutralizing rules than when applying neutralizing rules (i.e., higher accuracy on /s, z/ → [ʃ, ʒ] than on /t, d/ → [tʃ, dʒ] in Language A, and *vice versa* in Language B). The Novel test items were particularly relevant for testing the hypothesis. Whereas participants could succeed on Trained items by relying on their recognition of forms encountered during exposure, they had to learn a general rule in order to succeed on Novel items.

2.1. Method

2.1.1. Participants

Thirty native English speakers (17 females; mean age = 29) participated in Experiment 1. The experiment was conducted at University College London, and participants were recruited using a subject pool. Most participants were native speakers of Southern British English, though some spoke other varieties of English. They received a small amount of monetary compensation.

2.1.2. Materials

For the exposure phase, we created 48 CVCVC nonwords (henceforth, just 'words') as singular forms. For Language A, the consonants in all positions were chosen from a set of 12 phonemes: {p, b, k, g, f, v, t, d, s, z, tʃ, dʒ}. The 12 phonemes were assigned to the final consonant position an equal number of times, meaning that 4 words ended with each consonant. The consonants in initial and medial positions were roughly balanced in terms of frequency, distribution across positions, and vowel contexts. The following types of words were avoided: those similar to real English words, those with adjacent identical consonants (e.g., [takik]), and those with adjacent post-alveolar consonants (e.g., [zatʃadʒ]). Vowels were drawn from the set {i, a, u}. Each vowel was used approximately the same number of times and was evenly distributed across positions in the singular forms. Language B contained the same singular words as Language A, except that every [tʃ] and [dʒ] in Language A was replaced with [ʃ] and [ʒ], respectively, in Language B as per the different phoneme inventories of the two languages.

For each of the 48 singular forms, a corresponding plural form was created by adding the vowel /i/ to the end (i.e., CVCVC-i). The palatalization rules /t, d, s, z/ → [tʃ, dʒ, ʃ, ʒ] / __i were applied when the plural suffix was added to stems ending in /t, d, s, z/. The singular exposure words were chosen such that half of the words modified by the Neutralizing rules additionally resulted in homophony in the plural. For example, Language A (which had /t/ → [tʃ] as a Neutralizing rule) contained singular [tusut] and singular [tusutʃ], which were realized identically as [tusutʃi] in the plural, creating homophony. (Note that in Experiment 2, below, we manipulate amount of homophony as a variable to investigate how it affects the results.) In accordance with the palatalization rules, the sequences [ti], [di], [si], and [zi] never appeared in any stimuli.

Each singular word was randomly assigned to one of 48 pictures showing a singular object (e.g., a table). The corresponding plural word was assigned to a picture showing two or more of the same object (e.g., several tables). A singular/plural pair of words, coupled with their referential pictures, constituted a trial in the exposure phase. In all, there were 48 trials: 8 trials for alternating /t, d/ (Neutralizing in Language A), 8 trials for alternating /s, z/ (Neutralizing in Language B), 8 trials for singular words ending in non-alternating /tʃ, dʒ/ (Language A) or /ʃ, ʒ/ (Language B), and 24 filler trials for singular words ending in non-alternating /p, b, k, g, f, v/. To reiterate, the only difference between the two exposure languages was whether the words contained the non-alternating consonants /tʃ, dʒ/ or /ʃ, ʒ/.

The test phase consisted of 24 Trained items and 48 Novel items (72 in total). Two singular words ending in each consonant phoneme were chosen from the exposure phase to be used as the Trained items in the test phase. In addition, 48 Novel singular words were created in the same manner described above. The final consonants were distributed as in the exposure phase: four words ending in each of the alternating sounds, /t, d, s, z/, four words ending in the critical non-alternating sounds, /tʃ, dʒ/ (Language A) or /ʃ, ʒ/ (Language B), and four words ending in each of the filler sounds, /p, b, k, g, f, v/. Note that no minimal pairs were included in the Novel test items, so homophony was never created by the application of a Neutralizing rule in the Novel test phase (unlike in the exposure and Trained test phase).

For each singular form in the test phase, there was a correct plural form and an incorrect plural form. For trials in which palatalization was the correct option, the non-palatalized (unaltered) forms were used as incorrect options (e.g., singular [dazat], correct plural [dazatʃi], incorrect plural [dazati]). For filler trials, in which the non-palatalized (unaltered) form was the correct plural, an erroneously palatalized form was used as the incorrect plural, using the following correspondences: /p, b, k, g, f, v/ → [tʃ, dʒ, tʃ, dʒ, ʃ, ʒ]. Finally, for singular forms ending in /tʃ, dʒ/ or /ʃ, ʒ/, in which the final sound was already palatalized, we changed fricatives to affricates, or *vice versa*, to generate an incorrect plural form with a different palatalized sound (i.e., /tʃ, dʒ, ʃ, ʒ/ → [ʃ, ʒ, tʃ, dʒ]). Therefore, the plural options always included one non-changing form (e.g., [dazat]…[dazati]) and one changing form ending in one of the palatalized consonants, [tʃ, dʒ, ʃ, ʒ] (e.g., [dazat]…[dazatʃi]).

The stimuli were recorded by a male native speaker of English (phonetically trained) in a sound attenuated booth using a RØDE NT1-A condenser cardiod microphone and an audio interface recorder (RME Fireface UC). They were digitized at 44,100 Hz and 16 bits, and normalized in terms of amplitude. Stress was placed on the second syllable of all words.

2.1.3. Procedure

The experiment consisted of an exposure phase followed by two separate test phases, one for Trained items and one for Novel items. Before the exposure phase, participants were informed that they would be learning words in a foreign language. Participants were told to try to remember the words because they would be tested on them later. They were encouraged to repeat the words out loud to help them remember the words; we hoped this would make the phonological changes more salient and help participants maintain focus on the task. They were not given any information about the nature of the test or about the phonological patterns of interest.

The exposure phase included 48 self-paced trials. Each trial began with a singular picture appearing on the left side of the computer screen. The singular word for the item in the picture was played over the headphones after the picture had been shown for 1 s. After a pause of 2.5 s, the singular picture disappeared and the plural picture was displayed on the right side of the screen. The plural word for the picture was played after the picture had been shown for 1 s. The plural picture remained on the screen until participants decided to continue to the next trial by pressing the space bar. Presentation of the words was purely auditory; no orthographic forms were provided. Participants spent an average of about 15 min on the exposure phase.

The exposure phase was followed by two test phases; participants were always tested on Trained items first and Novel items second. In the Trained test phase, participants were tested on 24 Trained items chosen from the exposure phase. In the Novel test phase, they were tested on 48 Novel items. Before the Novel test phase, participants were told that they would be tested on new words belonging to the same language. The two test phases combined lasted approximately 15 min.

The test trials were identical to the exposure trials, except that when the plural picture was shown on the right side of the screen, participants heard two options for the plural word, the correct option and the incorrect option. The first option was played 1.5 s after the plural picture appeared, and the second option was played 1 s after the first option. Participants were forced to select one of the two options in order to initiate the next trial. They responded by pressing either the key marked ‘1’ for the first option that they heard or the key marked ‘2’ for the second option (the ‘f’ and ‘j’ keys, respectively, were marked as ‘1’ and ‘2’ for the responses). The order of the correct and incorrect options was counterbalanced; the correct option was played first for half of the stems ending in each consonant and the incorrect option was played first for the other half. The next trial started immediately after participants pressed a response key.

The full experiment lasted approximately 30 minutes. It was conducted in a sound attenuated room. The experimental software PsyScope (version X B77) was used to implement the experiment and record the responses.

2.2. Results

The data were analysed with mixed-effects logistic regression models (see Jaegar 2008), implemented in R (R Core Team, 2016) using the *lme4* package (Bates, Mächler, Bolker, & Walker, 2015). Following the recommendations of Barr, Levy, Scheepers, and Tily (2013), all models had a maximal random effects structure, including random intercepts for subjects and items as well as by-subject and by-item random slopes, unless the maximal model failed to converge. In the case of non-convergence, we removed as few random effects as possible to allow for convergence, and we started by removing random effects not related to our main variables of interest; for the sake of explicitness, we report the R code for the final models in footnotes below so that each model's random effects structure is clear.

As our aim was to investigate whether accuracy was greater for non-neutralizing rules than for neutralizing rules, we only include trials involving the alternating sounds (i.e., trials in which the correct answer involved the application of a rule) in the analyses below.[5] Figure

---

[5] Participants succeeded at learning which sounds should not be changed. Their mean accuracy on non-changing items was 73.3% for Trained items and 61.8% for Novel items, significantly higher than chance according to

1 shows the percentage of correct responses according to Item Type (Trained or Novel) and Rule Type (Neutralizing or Non-neutralizing). Recall that the Neutralizing rules in Language A were Non-neutralizing in Language B, and *vice versa*. Data from these two counterbalancing groups (Language A and Language B) are combined in Figure 1.

The initial model had fixed effects for Rule Type (Neutralizing or Non-neutralizing), Item Type (Trained or Novel), Counterbalancing Group (Language A or Language B), and all possible interactions. We then used backwards stepwise comparison to reach the final model, removing each fixed effect (except the intercept) one at a time and comparing the simpler model to the full model by way of a likelihood ratio test (using the *anova()* function; see Baayen 2008). Fixed effects that did not significantly improve model fit (at an alpha level of .05) were removed from the model. When multiple fixed effects could be removed, we started with the effect that had the highest p-value.

The fixed effects of the final model are presented in Table 2. We report p-values from two sources: (1) based on the Wald *z*, as reported in the model summary, and (2) based on a likelihood ratio test, as described in the previous paragraph.[6] There was a significant main effect of Item Type reflecting the fact that accuracy on Novel items (63.5%) was lower than accuracy on Trained items (72.1%) overall. The main effect of Rule Type was also significant. The interaction between Item Type and Rule Type did not significantly improve

---

intercept-only mixed-effects logit models (Trained items: $\beta = 1.29$, $z = 4.98$, $p < .01$; Novel items: $\beta = .61$, $z = 2.62$, $p < .01$). We do not consider these items further.

[6] In some cases, random effects needed to be removed to allow for the convergence of both the subset model and the superset model before the likelihood ratio test could be conducted, in order to maintain the same random effects structure in the two comparison models (see Barr et al., 2013). Though this can in principle affect the results, we see that the p-values from the likelihood ratio tests are highly consistent with those based on Wald *z* (which come from the model with the full random effects structure) for all of the analyses presented here.

model fit ($\chi^2(1) = .012$, $p = .91$) and thus did not warrant inclusion in the final model. The significant main effect of Rule Type, together with the lack of interaction effect, indicates that accuracy was lower for Neutralizing rules than for Non-neutralizing rules for both Trained items (Neutralizing: 67.5%; Non-neutralizing: 76.7%) and Novel items (Neutralizing: 57.9%; Non-neutralizing: 69.2%).

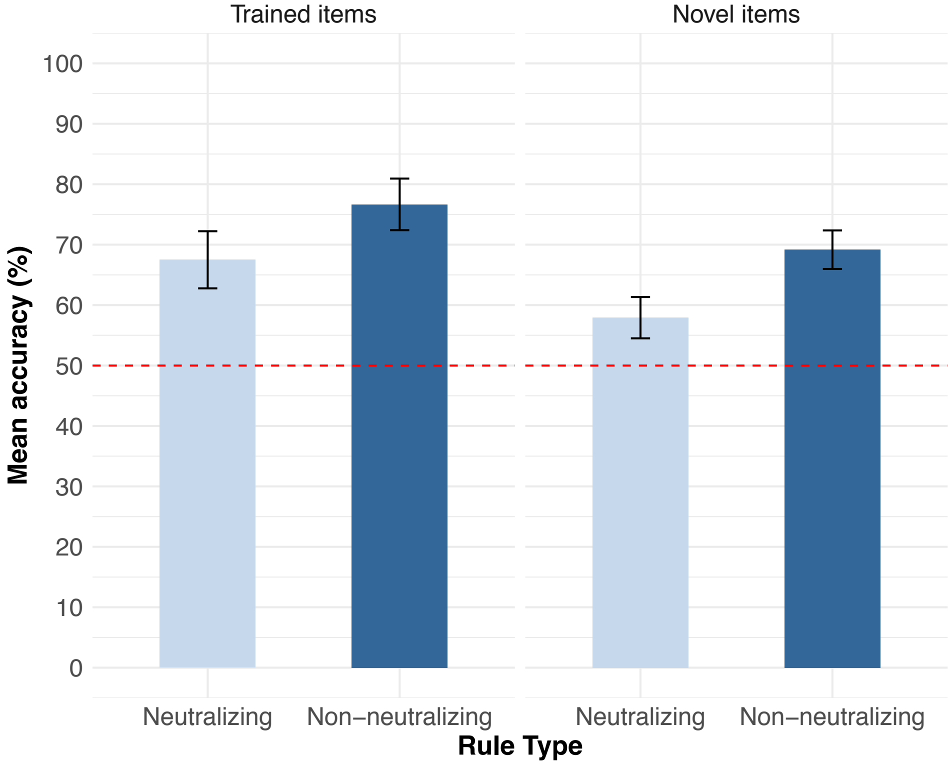

Figure 1. Percentage of test trials in which participants correctly chose the palatalized plural form in Experiment 1 by Item Type (Trained or Novel) and Rule Type (Neutralizing or Non-neutralizing). Errors bars show standard error of the mean. The dashed line indicates chance performance at 50%.

Table 2. Summary of the fixed effects for the final model in Experiment 1.[7]

Predictor	Estimate	Standard error	Wald *z*	p-value(Wald *z*)	p-value (likelihood ratio)
Intercept	1.03	.29	3.60	< .01	—
Rule Type = *Neutralizing*	– .61	.19	–3.25	< .01	< .01
Item Type = *Novel*	– .45	.18	–2.46	.01	.01
Counterbalancing Group = *Language A*	.76	.30	2.58	.01	.01

There was additionally a significant main effect of Counterbalancing Group, showing that participants who learned Language A (in which /t, d/ → [tʃ, dʒ] were the neutralizing rules) had an overall higher accuracy than those who learned Language B. The interaction between Rule Type and Counterbalancing Group, however, was not significant ($\chi^2(1) = .002$, $p = .97$), indicating that accuracy for Neutralizing rules was lower for participants in both counterbalancing groups. This can be seen in Figure 2.

---

[7] R code for model: glmer(Accuracy ~ RuleType + ItemType + CounterbalancingGroup + (1+RuleType|Subject) + (1|Item), data=data, family=binomial). The model failed to converge with by-subject random slopes for both Rule Type and Item Type, even when the correlations were removed; because our primary effect of interest was Rule Type, we removed the by-subject random slopes for Item Type to allow for convergence.

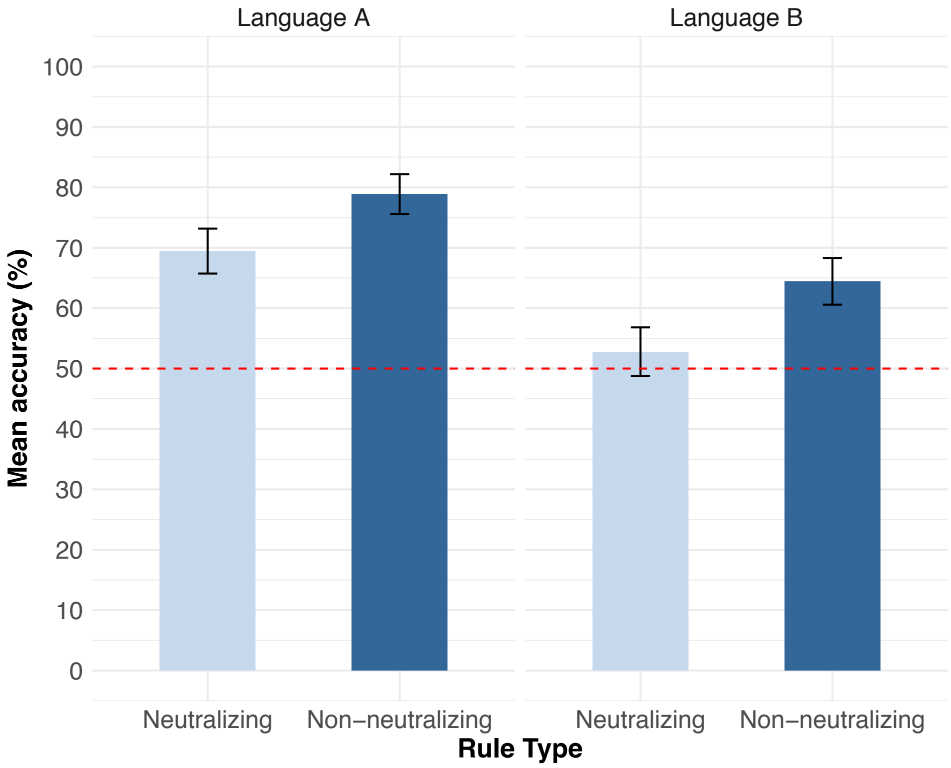

Figure 2. Percentage of test trials in which participants correctly chose the palatalized plural form in Experiment 1 by Counterbalancing Group (Language A or Language B) and Rule Type (Neutralizing or Non-neutralizing). Errors bars show standard error of the mean. The dashed line indicates chance performance at 50%.

Finally, in order to test whether participants actually learned the phonological rules and generalized them to novel items, we compared participants' accuracy to chance performance. We first divided the data into four sub-groups: (1) Trained items, Neutralizing rules; (2) Trained items, Non-neutralizing rules; (3) Novel items, Neutralizing rules; and (4) Novel items, Non-neutralizing rules. Note that these are the four groups of data represented in Figure 1. We then ran an intercept-only mixed logit model (i.e., one in which the only fixed effect was an intercept) for each subset. A significant positive intercept (in the intercept-only model) indicates that accuracy for that subset of data was significantly higher than the chance level of 50% (which would result in an intercept estimate of 0). Random intercepts for subject and item were also included in each model, with the exception of the model for Trained items and Non-neutralizing rules, which would only converge when the random intercepts for subjects were removed. As before, we report two p-values, one based on Wald *z* and one based on likelihood ratio tests comparing the intercept-only model to a simpler model without the intercept (i.e., a model with only random effects). A summary of the intercepts in the four intercept-only models is given in Table 3.

The intercept-only models show that accuracy on Trained items was significantly higher than chance for both Neutralizing rules and Non-neutralizing rules. For Novel items, accuracy was also significantly higher than chance for Non-neutralizing rules; however, the intercept for Neutralizing rules did not quite reach significance.

Table 3. Summary of the intercepts in the intercept-only models for Experiment 1.

Item Type	Rule Type	Estimate	Standard error	Wald z	p-value (Wald z)	p-value (likelihood ratio)
Trained	Neutralizing	.79	.24	3.29	< .01	< .01
	Non-neutralizing	1.19	.22	5.51	< .01	< .01
Novel	Neutralizing	.35	.19	1.87	.06	.07
	Non-neutralizing	.99	.26	3.81	< .01	< .01

2.3 Discussion

To summarize the results of Experiment 1, we found that participants learned the Non-neutralizing rules more easily than the Neutralizing rules. Compared to chance level, participants learned to apply both the Neutralizing rules and Non-neutralizing rules to Trained items. However, only the Non-neutralizing rules were applied to Novel forms at a greater-than-chance level. This suggests participants were able to recognize some of the

Trained items in both conditions, but they only (robustly) learned a general rule if the rule was Non-neutralizing. Moreover, even though participants performed better than chance on Trained items for both types of rules, their accuracy was significantly lower for Neutralizing rules than for Non-neutralizing rules.

Given the design of the experiment, the difference in learnability observed between Neutralizing and Non-neutralizing rules cannot be ascribed to any property of the rules themselves, other than their Neutralizing or Non-neutralizing status. The two counterbalancing groups learned minimally different languages. Both groups learned the same set of rules, but we manipulated which rules were Neutralizing or Non-neutralizing by making a minimal adjustment to the phoneme inventory of the language. The rules that were Neutralizing in Language A were Non-neutralizing in Language B, and *vice versa*. If something inherent to one of the rules made it harder to learn (e.g., due to its phonetic naturalness or complexity), then it would have resulted in reduced accuracy for both groups, whether the rule was Neutralizing or Non-neutralizing, resulting in an interaction between Rule Type and Counterbalancing Group. Instead, we saw that the *very same rule* (among an identical set of rules) was harder to learn when it was neutralizing and easier to learn when it was non-neutralizing.

There was a main effect of Counterbalancing Group: participants who learned Language A had higher accuracy overall than participants who learned Language B. The same set of rules were included in both languages, so we see no principled reason why Language A should have been easier than Language B overall. This disparity is most likely the result of individual differences in the groups of participants who learned the two languages (e.g., their motivation) despite their random assignment to conditions. Again, because the design was counterbalanced and there was no interaction with Rule Type, this difference in overall accuracy between the groups does not confound the results. Despite the difference in overall mean accuracy, Neutralizing rules were learned less well by both groups of participants.

Overall, the results of Experiment 1 suggest that learners have a bias against neutralization when learning phonological rules. But what role does homophony avoidance play in this effect? Is the neutralization avoidance effect driven by a bias against the creation of *actual* homophones, or is just the *possibility* of homophony (i.e. phonological neutralization, but no observed lexical neutralization) sufficient for triggering the bias? Experiment 1 was not designed to investigate these questions because the amount of homophony was fixed; the Neutralizing rules in Experiment 1 resulted in homophony in 50% of the forms where they applied. In Experiment 2, we investigate this question directly by manipulating whether or not the Neutralizing rules create homophony during exposure.

3. Experiment 2

Experiment 2 was designed to test whether the bias against neutralizing rules observed in Experiment 1 was driven by homophony avoidance. Specifically, we test whether the reduction in learnability that occurs when a rule is phonologically neutralizing depends on whether or not the rule creates *observed* cases of homophony. The design of Experiment 2 was identical to that of Experiment 1, except for the added variable for homophony.

Participants in Experiment 2 were randomly assigned to one of two conditions, Homophony or No Homophony, in addition to the two Counterbalancing Groups used in Experiment 1, Language A or Language B. Participants learned the same four rules from Experiment 1, [t, d, s, z] ~ [tʃ, dʒ, ʃ, ʒ]. As in Experiment 1, the /t, d/ → [tʃ, dʒ] rules were Neutralizing in Language A and the [s, z] → [ʃ, ʒ] rules were Neutralizing in Language B. In the Homophony condition, Neutralizing rules always created homophones in the plural by neutralizing a pair of distinct singular words. For example, in Language A, which contained neutralizing /t/ → [tʃ], singular [tusut] and singular [tusutʃ] were realized identically as [tusutʃi] in the plural. In the No Homophony condition, by contrast, Neutralizing rules did not create explicit homophones. In this case, Language A likewise contained singular [tusut] and plural [tusutʃi], but participants did not encounter singular [tusutʃ] during exposure; instead, they encountered a different singular word ending in [tʃ]. To summarize, in Experiment 1 50% of the exposure trials demonstrating a Neutralizing rule also involved lexical neutralization (i.e., creation of homophones) for all participants. In Experiment 2, Neutralizing rules created homophones 100% of the time in the Homophony condition, and 0% of the time in the No Homophony condition.

If homophony avoidance plays a role in the neutralization avoidance effect seen in Experiment 1, then the difference in accuracy between Neutralizing rules and Non-neutralizing rules should be greater in the Homophony condition than in the No Homophony condition. However, if homophony avoidance is not relevant when learning phonological patterns, then we should see a comparable learning deficit for Neutralizing rules in both the Homophony and No Homophony conditions.

3.1. Method

3.1.1. Participants

Forty native English speakers (27 females; mean age = 23) participated in Experiment 2. The recruitment process was identical to the process in Experiment 1. Participants received a small amount of monetary compensation.

3.1.2. Materials

The stimuli were designed and recorded following the same procedure used for Experiment 1. Most words were identical to the stimuli in Experiment 1, except for the singular exposure words ending in /t, d, tʃ, dʒ/ in Language A and /s, z, ʃ, ʒ/ in Language B, which had to be altered to implement the Homophony variable. All stimuli were recorded anew for Experiment 2 by the same speaker who recorded for Experiment 1.

In the Homophony condition of Language A, each singular word ending in /t/ or /d/ was paired with a minimally different word ending in /tʃ/ or /dʒ/, respectively, such that their plurals would be homophonous; for instance, [tusut] was paired with [tusutʃ], which both became [tusutʃi] in the plural. In total, there were four minimal pairs ending in /t/ and /tʃ/, and four minimal pairs ending in /d/ and /dʒ/. In Language B, where the Neutralizing rules were /s, z/ → [ʃ, ʒ], there were instead four minimal pairs for /s/ and /ʃ/ and four minimal pairs for /z/ and /ʒ/. In the No Homophony condition, no minimal pairs were included in the set of singular words, so each singular word had a unique plural form. The Novel test items were identical in the Homophony and No Homophony condition. Note that there were no minimal pairs in the Novel test items, so in the Homophony condition, homophony was created in the exposure and Trained test phases *only*, not in the Novel test phase.

3.1.3. Procedure

The procedure of Experiment 2 was identical to that of Experiment 1.

3.2. Results

The data were analysed using mixed-effects logistic regression models following the same procedure described for Experiment 1. Once again, we only include test trials involving the alternating sounds (i.e., those in which the correct response was the changing plural option) in the analyses below.[8] We analyze Trained items and Novel items separately.

3.2.1. Trained items

Figure 3 shows the accuracy for Trained items according to Rule Type and Homophony condition. The initial mixed logit model contained fixed effects for Rule Type (Neutralizing or Non-neutralizing), Homophony (Homophony or No Homophony), Counterbalancing

---

[8] As in Experiment 1, participants succeeded on the non-alternating items. Their mean accuracy was 74.8% for Trained items and 66.1% for Novel items, which was significantly higher than chance according to intercept-only mixed logit models (Trained items: $\beta = 1.43$, $z = 5.95$, $p < .01$; Novel items: $\beta = .77$, $z = 4.26$, $p < .01$). Accuracy on non-alternating items was similar across Homophony conditions for both Trained items (Homophony: 77.2%; No Homophony: 72.5%) and Novel items (Homophony: 67.5%; No Homophony: 64.7%). A mixed logit model for non-alternating items showed that the main effect of Homophony and all interactions involving Homophony were non-significant (all $p > .1$).

Group (Language A or B), and all possible interactions. The fixed effects of the final model (determined from model comparison using likelihood ratio tests, as described in section 2.2) are given in Table 4. There was a significant main effect of Homophony, reflecting the fact that overall accuracy on Trained items was higher in the No Homophony condition (78.1%) than in the Homophony condition (66.3%). There was also a significant main effect of Counterbalancing Group; overall accuracy on Trained items was higher for participants who learned Language B (78.1%) than for participants who learned Language A (66.3%).[9] The main effect of Rule Type, as well as the interaction effects, failed to significantly improve the model fit according to likelihood ratio tests (all $p > .1$).

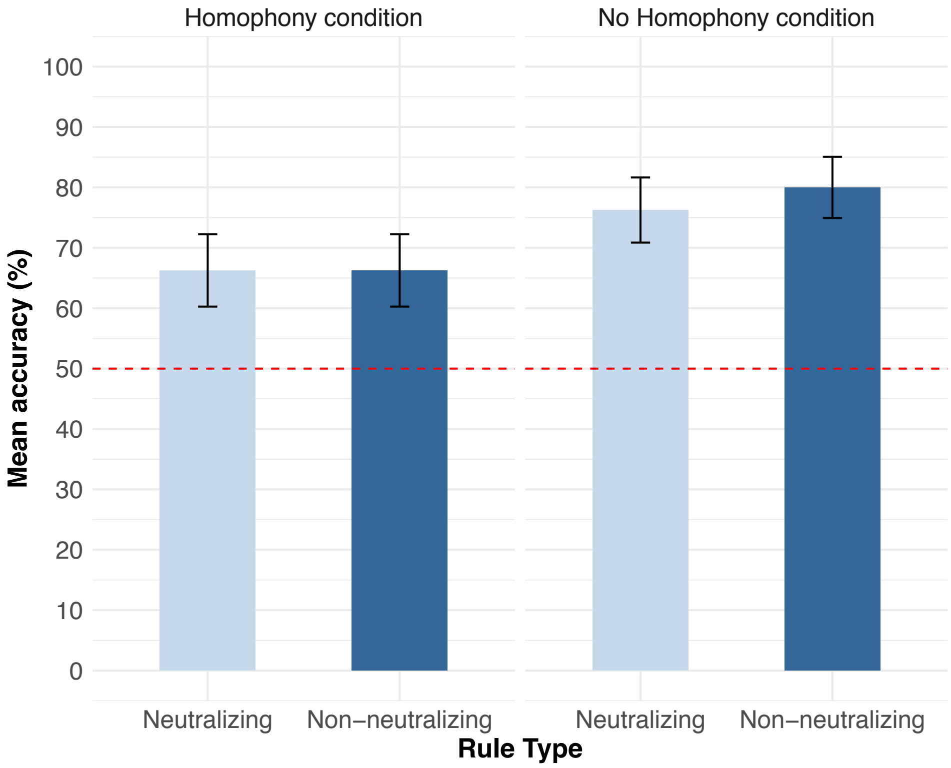

Figure 3. Percentage of Trained trials in which participants correctly chose the palatalized plural form in Experiment 2 by Homophony (Homophony condition or No Homophony condition) and Rule Type (Neutralizing or Non-neutralizing). Errors bars show standard error of the mean. The dashed line indicates

---

[9] The identical mean accuracies for these two comparisons is a coincidence, not a typographical mistake.

chance performance at 50%.

Table 4. Summary of the fixed effects for the final model for Trained items in Experiment 2.[10]

Predictor	Estimate	Standard error	Wald *z*	p-value (Wald *z*)	p-value (likelihood ratio)
Intercept	1.75	.32	5.54	< .01	—
Homophony = *Homophony*	–.67	.31	–2.14	.03	.03
Counterbalancing Group = *Language A*	–.66	.33	–2.03	.04	.04

3.2.2. Novel items

Figure 4 shows the accuracy for Novel items according to Rule Type and Homophony condition. The initial mixed logit model contained fixed effects for Rule Type (Neutralizing or Non-neutralizing), Homophony (Homophony or No Homophony), Counterbalancing Group (Language A or B), and all possible interactions. The fixed effects of the final model are given in Table 5. The main effects for Rule Type and Homophony were both non-significant. However, there was a significant Rule Type by Homophony interaction, reflecting the fact that accuracy on Neutralizing rules was lower than accuracy on Non-neutralizing rules, but only in the Homophony condition (Neutralizing: 47.5%; Non-neutralizing: 69.4%). In the No Homophony condition, accuracy was comparable for Neutralizing rules (66.9%) and Non-neutralizing rules (64.4%). Including the main effect of Counterbalancing Group, with or without its associated interaction effects, failed to significantly improve the model's fit according to likelihood ratio tests (all $p > .1$).

---

[10] R code for model: glmer(Accuracy ~ Homophony + CounterbalancingGroup + (1+RuleType|Subject) + (1+CounterbalancingGroup|Item), data=data, family=binomial).

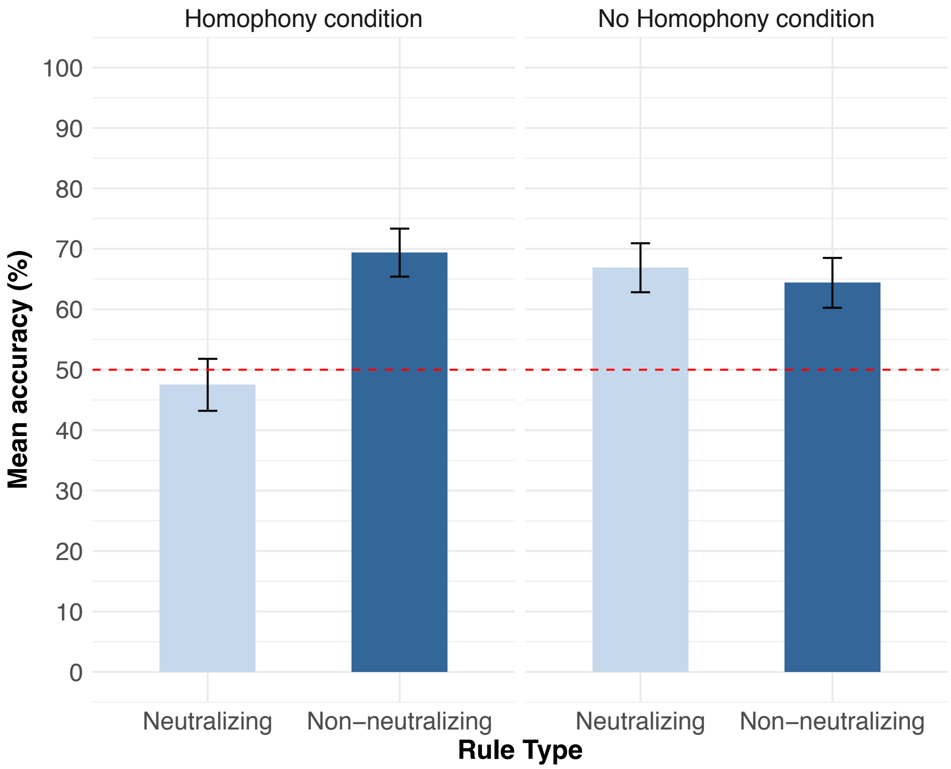

Figure 4. Percentage of Novel trials in which participants correctly chose the palatalized plural form in Experiment 2 by Homophony (Homophony condition or No Homophony condition) and Rule Type (Neutralizing or Non-neutralizing). Error bars show standard error of the mean. The dashed line indicates chance performance at 50%.

Table 5. Summary of the fixed effects for the final model, for Novel items in Experiment 2.[11]

Predictor	Estimate	Standard error	Wald *z*	p-value (Wald *z*)	p-value (likelihood ratio)
Intercept	.76	.32	2.38	.02	—
Rule Type = *Neutralizing*	.06	.30	.20	.84	—
Homophony = *Homophony*	.22	.43	.51	.61	—
Rule Type*Homophony = *Neut. & Homophony*	–1.17	.38	–3.08	< .01	< .01

---

[11] R code for model: glmer(Accuracy ~ RuleType*Homophony + (1+RuleType|Subject) + (1+CounterbalancingGroup|Item), data=data, family=binomial). Note that the non-significant main effects were included in the model due to the significant interaction between them.

3.2.3. Comparison to chance performance

To compare participants' accuracy to chance performance, we implemented intercept-only models as described for Experiment 1. A summary of the intercept-only models is given in Table 6. Overall, these models show that accuracy was significantly greater than chance in all conditions for Trained items. For Novel items, accuracy was not significantly different from chance for Neutralizing alternations in the Homophony condition (numerically, it was slightly below chance at 47.5%), but accuracy was significantly greater than chance in each of the other conditions.

Table 6. Summary of the intercepts in the intercept-only model for Experiment 2.[12]

Item Type	Subset	Estimate	Standard error	Wald z	p-value (Wald z)	p-value (likelihood ratio)
Trained	Homophony Neutralizing	.76	.32	2.42	.02	.01
	Homophony Non-neutralizing†	.79	.34	2.35	.02	.02
	No Homophony Neutralizing	1.24	.35	3.57	< .01	< .01
	No Homophony Non-neutralizing	1.62	.46	3.54	< .01	< .01
Novel	Homophony Neutralizing	– .11	.23	– .48	.63	.63
	Homophony Non-neutralizing	1.07	.35	3.05	< .01	< .01
	No Homophony Neutralizing†	.84	.29	2.95	< .01	< .01
	No Homophony Non-neutralizing	.72	.30	2.43	.02	.02

---

[12] The intercept-only models included random intercepts for subjects and items, except that the item intercepts had to be removed from the models marked with '†' to allow for convergence.

3.3. Discussion

The results from Experiment 2 showed that Neutralizing rules were harder to learn than Non-neutralizing rules, but only when the Neutralizing rules created homophony between lexical items observed during exposure. In the Homophony condition (and only in the Homophony condition), participants failed to generalize the Neutralizing rules to the Novel test items; in fact, their chance-level accuracy indicates that they were unable to acquire the Neutralizing rules at all in the Homophony condition. Though participants succeeded in applying the Neutralizing rules to the Trained items (indeed, their accuracy on Trained items was comparable for Neutralizing and Non-neutralizing rules in Exp. 2), this most likely stemmed from their ability to recognize plural words that they encountered during exposure. However, to succeed in extending the generalizations from the exposure items to the Novel test items, participants would have needed to learn a general phonological rule. As in Experiment 1, the counterbalancing measures taken (i.e., two exposure languages with the same rules but minimally different phoneme inventories) ensure that the effects observed in the experiment were due to the neutralizing status of the rules and homophony creation, rather than some inherent property of the rules themselves.

Overall, the results suggest that the neutralization avoidance effect observed in Experiment 1 was primarily driven by homophony avoidance, rather than an avoidance of phonological neutralization *per se*. Though the Neutralizing rules in both conditions involved phonological neutralization, only in the Homophony condition did the Neutralizing rules create homophony. Moreover, it is important to emphasize that the homophony in the

Homophony condition only occurred during the exposure phase (when participants were learning) and the Trained test phase (where participants could rely on recognition). The Novel test phase did *not* include any pairs of words that would be homophonous in the plural (the Novel items were in fact identical for the Homophony and No Homophony conditions). Therefore, it is not the case that participants actually learned the Neutralizing rules, but failed to apply them due to homophony during the Novel test phase. Rather, the results indicate that homophony avoidance interfered with participants' ability to learn the rules in the first place.

4. General discussion

Using two artificial language learning experiments, this study investigated whether the learnability of phonological rules is affected by their neutralizing status, and if so, to what extent this effect is driven by homophony avoidance. Experiment 1 established that neutralizing phonological rules (that created homophony) were less likely to be learned than identical non-neutralizing rules. Experiment 2 showed that neutralizing rules were only harder to learn when they produced homophony between lexical items encountered during learning. Together, the results indicate that learners have a bias against lexical neutralization, or homophony creation, and that this bias affects their learning of phonological rules. The results also suggest that mechanisms underlying the language learning process itself could serve as a source of homophony avoidance in language change.

In the following sections, we first consider, and reject, an account of the results based solely on phonological distributions (section 4.1). We then discuss in more detail the study's implications for models of grammar (section 4.2) and language change (section 4.3). Finally, we consider the relationship between neutralization avoidance and product-oriented learning, which in many cases make conflicting predictions (section 4.4).

4.1. Phonological distributions and neutralization

There is little doubt that tracking the distribution of speech sounds in the input plays a role in learning phonological rules. Pairs of sounds that are derived by rule from a single phoneme (‘allophones’) canonically have complementary distributions, meaning that the two sounds never occur in the same phonological context. By contrast, pairs of distinct phonemes canonically have overlapping distributions, meaning they do occur in the same phonological contexts (Hayes, 2009). By tracking the contexts of speech sounds, learners could exploit this distributional information to determine the phonological rules of their language (Peperkamp, Le Calvez, Nadal, & Dupoux, 2006). Could the experimental results be explained purely on the basis of phonological distributions? While there is indeed a plausible distributional account of the results of Experiment 1, detailed below, it cannot explain the difference between the Homophony and No Homophony conditions found in Experiment 2.

The account for Experiment 1 could work as follows. Neutralizing phonological rules create a more complex distributional pattern than non-neutralizing rules. This can be demonstrated using Language A as an example. Distinct phonemes that are not modified by any rules, such as /p/ and /tʃ/ in Language A, result in *fully overlapping distributions*; the sounds [p] and [tʃ] occur freely before all vowels. On the other hand, Non-neutralizing rules result in *complementary distributions*. For instance, the /s/ → [ʃ] rule is non-neutralizing in Language A, and the sounds [s] and [ʃ] are thus in complementary distribution: [ʃ] occurs only before the vowel [i] and [s] occurs only before vowels [a, u]. Neutralizing rules, however, result in distributions that are neither complementary nor fully overlapping; they are *partially overlapping*. For instance, the /t/ → [tʃ] rule is neutralizing in Language A. The sound [t] occurs before every vowel *except* [i]. The sound [tʃ] occurs before [i], but because /tʃ/ is itself a distinct phoneme, [tʃ] occurs before all other vowels as well. The only evidence that [tʃ] and [t] do not have fully overlapping distributions, which would make them distinct phonemes, is the absence of the sequence [ti]. By contrast, all three pre-vocalic contexts provide evidence that [ʃ] and [s] have different distributions because the sequences [si], [ʃa], and [ʃu] are all absent. These distributional differences are represented graphically in Figure 5.

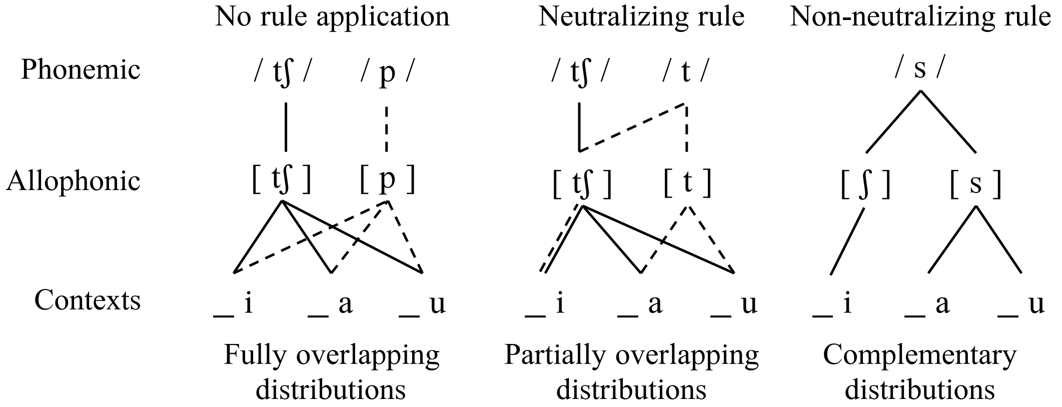

Figure 5. Sample distributional patterns from Language A resulting from the application of no rule, a neutralizing rule, and a non-neutralizing rule.

The fact that neutralizing rules result in partially overlapping distributions, whereas non-neutralizing rules result in complementary distributions, could provide a plausible account for why neutralizing rules were harder to learn in Experiment 1. Consistent with this idea, Peperkamp et al. (2006a) implemented a model capable of learning phonological rules by searching for complementary distributions; however, neutralizing rules were beyond the scope of the model due to the partially overlapping distributions that they create. Calamaro and Jarosz (2015) expanded the model, allowing it to measure distributional differences within specific contexts rather than collapsing across all contexts, which allowed the model to learn both neutralizing and non-neutralizing rules.[13] Calamaro and Jarosz do not explicitly address whether the model predicts a learning deficit for neutralizing rules compared to non-neutralizing rules. However, we can imagine an architecture in which more evidence is required for learners to treat a gap in a single context as meaningful (as with neutralizing rules) compared to gaps in many contexts (as with non-neutralizing rules).

Nevertheless, a purely distributional account cannot fully explain the results of the current study. The phonological distributions were identical in the Homophony and No Homophony conditions of Experiment 2, yet we saw a difference between neutralizing and non-neutralizing rules in the Homophony condition but not in the No Homophony condition. This suggests that although phonological distributions could have played a supporting role, the primary reason for the bias against neutralizing alternations in this study was homophony avoidance.

---

[13] Calamaro and Jarosz also allowed the model to learn rules at the level of phonological features rather than at the segmental level.

4.2. Implications for models of grammar and phonological learning

In order for homophony avoidance to have played a role in Experiment 2, learners must have been tracking the homophony that they encountered during exposure. Notably, we saw reduced accuracy for homophony-creating rules in the novel test items, but there was no actual homophony presented during the novel test phase. Homophonous forms were presented during the exposure phase, when participants were learning the rules. This means that homophony avoidance played a role at the time of learning, not (merely) at the time of application. Even though participants could recognize the trained items that they encountered (as evidenced by their performance on the trained test items), homophonous forms hindered their ability to induce a general phonological rule, which left them unable to apply the generalization to novel singular words. It is noteworthy that homophony between existing lexical items—actual homophony, not potential homophony—affected learners' ability to acquire phonological rules. This suggests a high degree of integration between lexical learning and phonological learning.

These findings have a natural explanation in exemplar-based models of phonology (e.g., Hay, 2000; Pierrehumbert, 2001, 2002; Bybee, 2002; Ernestus & Baayen, 2003; Wedel, 2004, 2006, 2007; Ussishkin & Wedel, 2009), which posit that the lexicon, the phonology, and the phonetics are highly integrated. In an exemplar model, the homophony avoidance effect could be modelled as lexical competition (e.g., see Wedel, 2004, 2006, 2012; Blevins & Wedel, 2009) with feedback throughout the system. Schematically, the model could work as follows. Competition between two singular words would interfere with setting up a two-to-one association between two distinct singular words and a shared plural form. Crucially, the weakened link between the alternating singular and plural forms must inhibit the learner's ability to extract a general phonological rule from the data. To take an example from the experiments, learners exposed to singular [tusut] and singular [tusutʃ] would set up distinct, competing lexical categories for the two words. When faced with [tusutʃi] as the plural of [tusut], competing [tusutʃ] ~ [tusutʃi] would inhibit the connection between [tusut] and [tusutʃi], which would in turn inhibit the formation of a general phonological rule relating [t] and [tʃ]. The connection between [tusutʃ] ~ [tusutʃi] would be independently bolstered (relative to [tusut] ~ [tusutʃi]) because no phonological change is required. Previous experiments show that learners are biased towards paradigm uniformity; that is, in the absence of conflicting evidence, learners assume that stems have the same phonological form throughout the paradigm, for instance, in both the singular and the plural (Tessier, 2012; Do, 2013; White, 2013, 2014, 2017).[14] In the event that singular [tusutʃ] is never encountered, as in the No Homophony condition, lexical competition would not occur, and the connection between [tusut] and [tusutʃi] would not be inhibited. The pair would thus provide evidence for a general phonological rule relating [t] and [tʃ], which could then be applied to novel words.

It is possible that the bias observed in these experiments is an instantiation of a more general bias against two-to-one mappings. We know that lexical development is constrained

---

[14] This raises the question of how neutralization avoidance interacts with paradigm uniformity during learning. In the experiments reported here, one phoneme always merged into another phoneme (e.g., phonemes A and B neutralized to B). It would be interesting to compare this type of neutralization to a case where two phonemes were neutralized into a third sound (e.g., A and B neutralized to C), like the case of flapping in American English. We leave this issue for future research.

by the mutual exclusivity bias, whereby children as young as 17 months of age prefer one-to-one mappings between words and referents (Markman & Wachtel, 1988; Merriman & Bowman, 1989; Halberda, 2003). For instance, children will assume that a newly encountered word refers to a novel item rather than a familiar item, presumably to avoid mapping the same meaning to more than one word, or the same word to more than one meaning (Clark, 1988; Merriman & Bowman, 1989). Similarly, children from a wide range of ages (from 2;9 to 11;8) have difficulty interpreting the meaning of newly coined words from context when the new words are homophonous with existing words (Mazzocco, 1997). These effects, though in the lexical-semantic domain, appear strikingly similar to the homophony avoidance effect that we observed in the current study. Thus, the same general bias against two-to-one mappings might play a role both in the interaction between lexical learning and semantic learning, and in the interaction between lexical learning and phonological learning.

Our findings pose a challenge for traditional generative models of phonology, which have typically maintained a modular view of the lexicon and the phonology. There have been several proposals within Optimality Theory (Prince & Smolensky, 1993/2004) and Dispersion Theory (Flemming, 1995, 2004; Padgett, 2003) to account for neutralization avoidance or contrast maintenance effects (Crosswhite, 1999; Łubowicz, 2003; Padgett, 2003, 2009; Ito & Mester, 2004; Ní Chiosáin & Padgett, 2009;). However, many of these proposals have taken the position that neutralization avoidance is applicable at the level of *possible* words, not *actual* words in the lexicon (Łubowicz, 2003; Padgett, 2003, 2009; Ní Chiosáin &

Padgett, 2009); this restriction is aligned with the traditional generative phonology view that the role of phonology is limited to describing the possible words of a language rather than the precise contents of the lexicon (e.g., Halle, 1962). In effect, this approach allows for neutralization avoidance at the phonological level (i.e., the prevention of phonemes being neutralized), but it precludes homophony at the lexical level from influencing phonological patterning. Other proposals (Crosswhite, 1999; Ito & Mester, 2004) have permitted homophony avoidance effects, but have limited them to paradigmatically related forms; for instance, homophony between the singular and plural forms of the same noun could be prevented, but not homophony between the plural forms of two distinct nouns.

The results of the current study are inconsistent with either approach. First, neutralizing rules were only more difficult to learn when they created homophony between observed lexical items. This suggests phonological learning is indeed sensitive to the contents of the lexicon, including the presence of homophony. Second, homophony avoidance was observed in the experiment between forms that were not paradigmatically related (i.e., the plural forms of two distinct nouns), suggesting that learners are biased to avoid homophony not only within paradigms. To account for the results, generative models would need to be expanded to include phonological constraints that are sensitive to homophony even for forms that are not paradigmatically related (e.g., see Ichimura, 2006).

4.3. Mechanisms driving homophony avoidance in language change

There are several mechanisms through which homophony avoidance could influence language change (for an overview, see Kaplan, 2011). Previous research has mainly focused on mechanisms external to the learning process in seeking to explain homophony avoidance in language change. For instance, Wedel (2004, 2006, 2012) and Blevins and Wedel (2009) emphasize the role of production and perception. They show that homophony avoidance will arise from a production-perception feedback loop in an exemplar model containing lexical competition and a bias that serves to enhance the contrast between competing words in production. Consistent with this idea, speakers tend to hyperarticulate words that are in dense neighborhoods, where ambiguous productions would be more likely to cause misidentification (Munson & Solomon, 2004; Baese & Goldrick, 2009; Scarborough, 2010). Wedel's (2004, 2006, 2012) model implements the bias as a reduction in the influence of exemplars that are ambiguous between two competing lexical categories. The model therefore places the locus of explanation for homophony avoidance in production and perception, not in the learning process itself.

The results from the current study, however, suggest that a learning bias could also play a role in promoting homophony avoidance—and thus maintaining lexical contrasts—during the course of language change. Even if gradual phonetic shifts created the conditions that could lead to the adoption of a new neutralizing phonological rule, the current study suggests that learners would be less likely to adopt the rule if it resulted in a large amount of homophony. These mechanisms are not mutually exclusive. Indeed, it is possible that learning biases and factors related to production and perception work together to limit homophony over the course of language change (e.g., see Moreton, 2008 for arguments that learning biases and production/perception both play a role in explaining language change and typology).

The current results diverge from the findings of several word learning experiments using the iterated learning paradigm (Kirby, Cornish, & Smith, 2008; Kirby, Tamariz, Cornish, & Smith, 2015), which have suggested that homophony avoidance arises from communicative pressures external to the learning process. In these experiments, participants learned novel labels for abstract pictures, and the output from one 'generation' of speakers was used as the input for the next group of participants, simulating language transmission across generations in the lab. Notably, when participants completed the learning task alone, they readily reduced the number of distinct labels (over several generations) such that multiple pictures were assigned the same label (Kirby et al., 2008). However, when a communicative component was added to the task, whereby participants were incentivized to ensure that a partner could accurately choose the correct picture based on the label provided, the languages were simplified in a more structured way that better maintained the contrasts between the pictures (Kirby et al., 2015). For instance, in one transmission chain, the language developed a set of suffixes to systematically mark a shape's fill pattern. Kirby et al. (2015) concluded that a bias to simplify the grammar is active during learning, whereas a bias to avoid ambiguity or loss of contrast ('expressivity') arises not from learning, but rather from communicative pressures external to the learning process. Indeed, the authors argue that learning alone tends to lead to highly degenerate languages. These studies suggest that homophony avoidance only occurs in tasks involving a communicative component (see also Kanwal, Smith, Culbertson, & Kirby, 2017).

In the present study, however, participants were not under any particular communicative pressure, as the task involved solitary learning with no direct motivation to maintain the distinctions between words. Moreover, as mentioned above, homophony was included when participants were learning (i.e., in the exposure phase), but not when participants were tested on novel words. Therefore, the homophony avoidance bias observed in the current study seems to be active during learning, rather than stemming solely from communicative pressures external to the learning process.

A potential explanation for the diverging results may lie in the distinction between homophony (one form with multiple distinct meanings) and polysemy (one form with multiple related meanings), which have been shown to have differential effects on language processing (e.g., Rodd, Gaskell, & Marslen-Wilson, 2002; Pylkkänen, Llinás, & Murphy, 2006; Simon, Lewis, & Marantz, 2012). In the label-learning tasks (Kirby et al., 2008; Kirby et al., 2015), participants were required to learn and produce novel labels for a large number of abstract pictures with a significant amount of semantic overlap (e.g., a limited number of abstract shapes, crossed with a limited number of textures). Participants who could not recall the correct labels would be forced either to make up an arbitrary label, or to adopt one of the other labels that they could remember. Merging labels in such a task could be considered akin to creating polysemy by generalizing labels to include similar objects. Even in natural language, it is common to generalize labels to refer to semantically related objects, especially when there is a failure of lexical access. For example, it is natural to extend the label ‘chicken’ to refer to a guinea fowl, or any chicken-like bird, especially if one cannot recall the bird's real name. In this case, 'chicken' is polysemous as it could refer to the specific animal that is a chicken, or it could refer to any small bird that is similar to a chicken.

On the other hand, participants in the current study learned novel names for everyday objects that had little semantic overlap. Neutralized words therefore resulted in homophony rather than polysemy. The nature of the task was also different. Participants in the current study were tested on their ability to apply phonological rules to novel words that were provided to them at test, so failure to recall the words themselves was not at issue in the current study. In sum, the diverging findings of these studies may stem from the distinction between homophony and polysemy, in combination with different task demands and the different types of learning involved in the tasks (label learning vs. phonological learning). Further research is needed to disentangle these differences.

Finally, the results of the current study are generally consistent with the functional load hypothesis (Martinet, 1952; Hockett, 1967; King, 1967; Wedel et al., 2013) as applied to learners: neutralizing rules are only dispreferred if they create homophony, which increases ambiguity. Taking the hypothesis one step further, the learnability of a neutralizing rule is predicted to vary according to the amount of homophony that it creates. Though we only tested three levels of homophony creation across the two experiments (0%, 50%, and 100%), the results are broadly consistent with this more nuanced hypothesis. As shown in Table 7, there was a gradual decrease in accuracy for Neutralizing rules in the Novel test phase as the amount of homophony increased; accuracy for Non-neutralizing rules remained relatively stable across the three conditions. As a post hoc exploratory analysis, we implemented a global mixed-effects logit model of the Novel test items across the two experiments (following the same procedure described for the analyses above) as a way to compare the three levels of homophony that we tested (0%, 50%, and 100% homophony, coded as a categorical factor). The final model revealed that accuracy for Neutralizing rules at the 0% Homophony level was significantly different from accuracy at the 100% level ($p < .01$), but the difference between the 0% and 50% levels was non-significant ($p = .15$). When we recoded 50% as the reference level, we found that the difference between the 50% and 100% levels was also non-significant ($p = .15$). Thus, while the results provide preliminary support for the more nuanced hypothesis of a graded homophony influence, only the most extreme comparison (0% vs. 100% homophony) reaches significance in this study. Future experiments could be designed to provide a more complete test of the graded hypothesis, for instance by investigating a greater range of homophony creation (e.g., 0%, 20%, 40%, 60%, 80%, 100%).

Table 7. Summary of accuracy (% correct) for Novel test items across Experiments 1 and 2.

	0% homophony (Exp. 2, No Homophony)	50% homophony (Exp. 1)	100% homophony (Exp. 2, Homophony)
Neutralizing rules	66.9	57.9	47.5
Non-neutralizing rules	64.4	69.2	69.4

4.4. Neutralization avoidance and product-oriented learning

At first glance, our results seem to diverge from those of Kapatsinski (2013), who also conducted an artificial language experiment involving palatalizing rules. In Kapatsinski's study, English-speaking participants learned velar palatalization, specifically /k/ → [tʃ] before the vowel [i], and were then tested to see whether the rule would be generalized to other voiceless stops (i.e., /t/ → [tʃ] and /p/ → [tʃ]). The most notable finding in relation to the current study is that when cases of /tʃ/ → [tʃ] were included in the training, which would cause the /t/ → [tʃ] to be neutralizing, participants were *more likely* to generalize to the /t/ → [tʃ] change. Kapatsinski interpreted the results as evidence that participants had learned a product-oriented schema (Bybee & Slobin, 1982), for instance, that [tʃi] was a particularly good ending for plurals regardless of the source. Under this account, even though pairs of singular [tʃ]-final words and plural [tʃi]-final words provided no positive evidence for any change, they reinforced the product-oriented schema that [tʃi] was a good plural ending, thus increasing the motivation to choose a [tʃi]-final plural for a [t]-final singular word.

In the current study, we found that including cases of non-changing /tʃ/ *reduced* the likelihood that participants would adopt the neutralizing /t/ → [tʃ] rule, consistent with a neutralization avoidance bias rather than product-oriented learning. However, there are crucial differences between the two studies that likely explain the different results. Most importantly, Kapatsinski did not include minimal pairs in the training data, meaning that no actual homophony was created by applying any of the changes. Our results suggest that potential homophony (as opposed to actual homophony) is not sufficient to generate a neutralization avoidance effect; thus we would not expect to find neutralization avoidance in Kapatsinski's study.

Though our study was not designed to test specifically for product-oriented learning, we conducted a post-hoc analysis of the filler items to see if there was any evidence for product-oriented learning in the experiment. Note that product-oriented schemas predict the opposite of the observed results for the target items, so the target items do not provide evidence for product-oriented learning, hence why we turn to the filler items. In Language A, where /tʃ/ and /dʒ/ were phonemes, there was more evidence for [tʃi] and [dʒi] plural schemas because more plural forms ended in these sequences; in Language B, where /ʃ/ and /ʒ/ were phonemes, there was more evidence for [ʃi] and [ʒi] plural schemas. We compared the error rates in Experiment 1 for filler stops /p, b, k, g/, whose incorrect plural option in the forced-choice task ended in [tʃi] and [dʒi], to errors rates for filler fricatives /f, v/, whose incorrect plural options ended in [ʃi] and [ʒi]. If participants were learning product-oriented schemas, they should have been more likely to change (incorrectly) filler stops in Language A (where plurals ending in [tʃi] and [dʒi] were more common) and filler fricatives in Language B (where plurals ending in [ʃi] and [ʒi] were more common). In fact, error rates on the filler sounds were almost identical across the two languages (Lang. A, stops: 33.9%; Lang. A, fricatives: 30.6%; Lang. B, stops: 33.6%; Lang. B, fricatives: 30.6%), suggesting that product-oriented learning played no role in our study.

Nevertheless, we think that learners do make product-oriented generalizations, at least under some learning conditions. Several differences between our study and Kapatsinski's study could explain why we failed to see such an effect in our study. First, Kapatsinski’s study involved learning a single type of change (/k/ → [tʃ]), whereas our study involved learning four distinct rules resulting in a wide variety of plural endings. Perhaps these conditions did not provide sufficient consistency for participants to develop robust product-oriented schemas after the limited exposure that they received. A second difference is the type of response required during the test phase: Kapatsinski’s study involved free oral production whereas the current experiments involved forced-choice. It is possible that participants are more likely to rely on product-oriented schemas when they must produce a free oral response. More research is needed to better understand the interaction between neutralization (or homophony) avoidance and product-oriented generalizations during phonological learning, particularly given that they often make opposing predictions.

5. Conclusions

This study demonstrated that learners have a bias against neutralizing phonological rules, but only if those rules create homophony between existing lexical items. We draw two main conclusions from these findings. First, the results indicate that there is a high degree of interaction between lexical learning and phonological learning, suggesting that the part of the grammar responsible for making phonological generalizations must be sensitive to the contents of the lexicon. Second, the findings suggest that a learning bias could serve as a mechanism that promotes homophony avoidance during language change, though other mechanisms likely also play a role.

6. Acknowledgments

For helpful discussion and comments, we would like to thank Wing Yee Chow, Gillian Gallagher, Vsevolod Kapatsinski, Andrew Nevins, Yasutada Sudo, Adam Ussishkin, and three anonymous reviewers, as well as audiences of the London Phonology Seminar, the 2015 Annual Meeting on Phonology (AMP), and the 2016 annual meeting of the LSA. We would also like to thank Andrew Clark for help in the lab. Any remaining errors are ours alone.

7. References

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Baer-Henney, D., & van de Vijver, R. (2012). On the role of substance, locality, and amount of exposure in the acquisition of morphophonemic alternations. *Laboratory Phonology*, *3*(2), 221–249.

Baese-Berk, M., & Goldrick, M. (2009). Mechanisms of interaction in speech production. *Language and Cognitive Processes*, *24*(4), 527–554.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using {lme4}. *Journal of Statistical Software*, *67*(1), 1–48.

Blevins, J., & Wedel, A. (2009). Inhibited sound change: An evolutionary approach to lexical competition. *Diachronica*, *26*(2), 143–183.

Bybee, J. (2002). Phonological evidence for exemplar storage of multiword sequences. *Studies in Second Language Acquisition*, *24*(2), 215–221.

Bybee, J. L., & Slobin, D. I. (1982). Rules and schemas in the development and use of the English past tense. *Language, 58*(2), 265–289.

Calamaro, S., & Jarosz, G. (2015). Learning general phonological rules from distributional information: A computational model. *Cognitive Science*, *39*(3), 647–666.

Carpenter, A. C. (2010). A naturalness bias in learning stress. *Phonology*, *27*(3), 345–392.

Chambers, K. E., Onishi, K. H., & Fisher, C. (2003). Infants learn phonotactic regularities from brief auditory experience. *Cognition*, *87*(2), B69--B77.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper & Row.

Clark, E. V. (1988). On the logic of contrast. *Journal of Child Language*, *15*(2), 317–335.

Cristià, A., & Seidl, A. (2008). Is infants' learning of sound patterns constrained by phonological features? *Language Learning and Development*, *4*(3), 203–227.

Crosswhite, K. (1999). Intra-paradigmatic homophony avoidance in two dialects of Slavic. *UCLA Working Papers in Linguistics*, *1*, 48–67.

Culbertson, J., & Newport, E. L. (2015). Harmonic biases in child learners: In support of language universals. *Cognition*, *139*, 71–82.

Culbertson, J., Smolensky, P., & Legendre, G. (2012). Learning biases predict a word order universal. *Cognition*, *122*(3), 306–329.

Do, Y. (2013). *Biased learning of phonological alternations*. Ph.D. dissertation, MIT.

Ernestus, M., & Baayen, R. H. (2003). Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language*, 5–38.

Fehér, O., Wonnacott, E., & Smith, K. (2016). Structural priming in artificial languages and the regularisation of unpredictable variation. *Journal of Memory and Language*, *91*, 158–180.

Finley, S. (2011). The privileged status of locality in consonant harmony. *Journal of Memory and Language*, *65*(1), 74–83.

Finley, S. (2017). Learning metathesis: Evidence for syllable structure constraints. *Journal of Memory and Language*, *92*, 142–157.

Finley, S., & Badecker, W. (2009). Artificial language learning and feature-based generalization. *Journal of Memory and Language*, *61*(3), 423–437.

Finley, S., & Badecker, W. (2012). Learning biases for vowel height harmony. *Journal of Cognitive Science*, *13*(3), 287–327.

Flemming, E. (1995). *Auditory Representations in Phonology*. Phd Dissertation, UCLA.

Flemming, E. (2004). Contrast and perceptual distinctiveness. In B. Hayes, R. Kirchner, & S. Donca (Eds.), *Phonetically-Based Phonology* (pp. 232–276). Cambridge: Cambridge University Press.

Gilliéron, J. (1918). *Généalogie des mots qui désignent l'abeille d'après l'Atlas linguistique de la France* (Vol. 225). Paris: É. Champion.

Gómez, R. L., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences*, *4*(5), 178–186.

Halberda, J. (2003). The development of a word-learning strategy. *Cognition*, *87*(1), B23–B34.

Halle, M. (1962). Phonology in generative grammar. *Word*, *18*(1–3), 54–72.

Hay, J. (2000). *Causes and Consequences of Word Structure*. Ph.D. dissertation, Northwestern University.

Hayes, B. (2009). *Introductory phonology*. Oxford: Wiley-Blackwell.

Herd, W., Jongman, A., & Sereno, J. (2010). An acoustic and perceptual analysis of /t/ and /d/ flaps in American English. *Journal of Phonetics*, *38*(4), 504–516.

Hockett, C. F. (1967). The quantification of functional load. *Word*, *23*(1–3), 300–320.

Ichimura, L. K. (2006). *Anti-homophony blocking and its productivity in transparadigmatic relations*. Ph.D. dissertation, University of Massachusetts Boston.

Ito, J., & Mester, A. (2004). Morphological contrast and merger: Ranuki in Japanese. *Journal of Japanese Linguistics*, *20*, 1–18.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446.

Jakobson, R. (1931). Prinzipien der historischen Phonologie. *Travaux Du Cercle Linguistique de Prague*, *4*, 246–267.

Kalish, M. L., Griffiths, T. L., & Lewandowsky, S. (2007). Iterated learning: Intergenerational knowledge transmission reveals inductive biases. *Psychonomic Bulletin & Review*, *14*(2), 288–294.

Kanwal, J., Smith, K., Culbertson, J., & Kirby, S. (2017). Zipf's law of abbreviation and the principle of least effort: Language users optimize a miniature lexicon for efficient communication. *Cognition, 165*, 45–52.

Kapatsinski, V. (2013). Conspiring to mean: Experimental and computational evidence for a usage-based harmonic approach to morphophonology. *Language*, *89*(1), 110–148.

Kaplan, A. (2011). How much homophony is normal? *Journal of Linguistics*, *47*(3), 631–671.

Kaplan, A. (2015). The evidence for homophony avoidance in language change: Reply to Sampson (2013). *Diachronica, 32*(2), 268–276.

King, R. D. (1967). Functional load and sound change. *Language*, 831–852.

Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences, 105*(31), 10681–10686.

Kirby, S., Smith, K., & Brighton, H. (2004). From UG to universals: Linguistic adaptation through iterated learning. *Studies in Language. International Journal Sponsored by the Foundation "Foundations of Language," 28*(3), 587–607.

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition, 141*, 87–102.

Łubowicz, A. (2003). *Contrast preservation in phonological mappings*. Ph.D. dissertation, University of Massachusetts Amherst.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*(2), 121–157.

Martinet, A. (1952). Function, structure, and sound change. *Word*, *8*(1), 1–32.

Mathesius, V. (1931). Zum Problem der Belastungs- und kombinationsfähigkeit der Phoneme. *Travaux Du Cercle Linguistique de Prague*, *4*, 148–152.

Mazzocco, M. M. M. (1997). Children's interpretations of homonyms: a developmental study. *Journal of Child Language*, *24*(2), 441–467.

Merriman, W. E., Bowman, L. L., & MacWhinney, B. (1989). The mutual exclusivity bias in children's word learning. *Monographs of the Society for Research in Child Development*, *54*(3–4), i-129.

Moreton, E. (2008). Analytic bias and phonological typology. *Phonology*, *25*(1), 83–127.

Moreton, E., & Pater, J. (2012a). Structure and substance in artificial-phonology learning, part II: Substance. *Language and Linguistics Compass*, *6*(11), 702–718.

Moreton, E., & Pater, J. (2012b). Structure and substance in artificial-phonology learning, Part I: Structure. *Language and Linguistics Compass*, *11*(6), 686–701.

Munson, B., & Solomon, N. P. (2004). The effect of phonological neighborhood density on vowel articulation. *Journal of Speech, Language, and Hearing Research*, *47*(5), 1048–1058.

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*(2), 127–162.

Ní Chiosáin, M., & Padgett, J. (2009). Contrast, comparison sets, and the perceptual space. In S. Parker (Ed.), *Phonological argumentation: Essays on evidence and motivation* (pp. 103–121). London: Equinox.

Onishi, K. H., Chambers, K. E., & Fisher, C. (2002). Learning phonotactic constraints from brief auditory experience. *Cognition*, *83*(1), B13--B23.

Padgett, J. (2003). Contrast and post-velar fronting in Russian. *Natural Language and Linguistic Theory*, *21*(1), 39–87.

Padgett, J. (2009). Systemic contrast and Catalan rhotics. *Linguistic Review*, *26*(4), 431–463.

Peperkamp, S., Le Calvez, R., Nadal, J.-P., & Dupoux, E. (2006a). The acquisition of allophonic rules: Statistical learning with linguistic constraints. *Cognition*, *101*(3), B31--B41.

Peperkamp, S., Skoruppa, K., & Dupoux, E. (2006b). The role of phonetic naturalness in phonological rule acquisition. In D. Bamman, T. Magnitskaia, & C. Zaller (Eds.), *Proceedings of the 30th annual Boston University Conference on Language Development* (pp. 464–475). Somerville, MA: Cascadilla Press.

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition, and contrast. In J. Bybee & P. Hopper (Eds.), *Frequency effects and the emergence of linguistic structure* (pp. 137–157). Amsterdam: John Benjamins.

Pierrehumbert, J. (2002). Word-specific phonetics. *Laboratory Phonology*, *VII*, 101–139.

Prince, A., & Smolensky, P. (1993/2004). *Optimality theory: Constraint interaction in generative grammar*. Cambridge: Blackwell (Published in 2004. First circulated in 1993).

Pylkkänen, L., Llinás, R., & Murphy, G. L. (2006). The representation of polysemy: MEG evidence. *Journal of Cognitive Neuroscience, 18*(1), 97–109.

R Core Team. (2016). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. Vienna, Austria. Retrieved from http://www.r-project.org/

Reali, F., & Griffiths, T. L. (2009). The evolution of frequency distributions: Relating regularization to inductive biases through iterated learning. *Cognition*, *111*(3), 317–328.

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

Saffran, J. R., & Thiessen, E. D. (2003). Pattern induction by infant language learners. *Developmental Psychology*, *39*(3), 484–494.

Sampson, G. (2013). A counterexample to homophony avoidance. *Diachronica*, *30*(4), 579–591.

Scarborough, R. (2010). Lexical and contextual predictability: Confluent effects on the production of vowels. *Laboratory Phonology*, *10*, 557–586.

Seidl, A., & Buckley, E. (2005). On the learning of arbitrary phonological rules. *Language Learning and Development*, *1*(3–4), 289–316.

Silverman, D. (2010). Neutralization and anti-homophony in Korean. *Journal of Linguistics*, *46*(2), 453–482.

Silverman, D. (2012). *Neutralization*. Cambridge: Cambridge University Press.

Simon, D. A., Lewis, G., & Marantz, A. (2012). Disambiguating form and lexical frequency effects in MEG responses using homonyms. *Language and Cognitive Processes, 27*(2), 275–287.

Skoruppa, K., & Peperkamp, S. (2011). Adaptation to Novel Accents: Feature-Based Learning of Context-Sensitive Phonological Regularities. *Cognitive Science*, *35*(2), 348–366.

Smith, K., & Wonnacott, E. (2010). Eliminating unpredictable variation through iterated learning. *Cognition*, *116*(3), 444–449.

Tessier, A.-M. (2012). Testing for OO-Faithfulness in the acquisition of consonant clusters. *Language Acquisition*, *19*(2), 144–173.

Ussishkin, A., & Wedel, A. (2009). Lexical access, effective contrast and patterns in the lexicon. *Perception in Phonology*, 267–292.

Wedel, A. (2004). Category competition drives contrast maintenance within an exemplar-based production/perception loop. In J. A. Goldsmith & R. Wicentowski (Eds.), *Proceedings of the 7th Meeting of the ACL Special Interest Group in Computational Phonology* (pp. 1–10). Barcelona: Association for Computational Linguistics.

Wedel, A. (2006). Exemplar models, evolution and language change. *The Linguistic Review*, *23*(3), 247–274.

Wedel, A. (2007). Feedback and regularity in the lexicon. *Phonology*, *24*(1), 147–185.

Wedel, A. (2012). Lexical contrast maintenance and the organization of sublexical contrast systems. *Language and Cognition*, *4*(4), 319–355.

Wedel, A., Kaplan, A., & Jackson, S. (2013). High functional load inhibits phonological contrast loss: A corpus study. *Cognition*, *128*(2), 179–186.

White, J. (2013). *Bias in phonological learning: evidence from saltation*. Ph.D. dissertation, UCLA.

White, J. (2014). Evidence for a learning bias against saltatory phonological alternations. *Cognition*, *130*(1), 96–115.

White, J. (2017). Accounting for the learnability of saltation in phonological theory: A maximum entropy model with a P-map bias. *Language*, *93*(1), 1–36.

White, J., & Sundara, M. (2014). Biased generalization of newly learned phonological alternations by 12-month-old infants. *Cognition*, *133*(1), 85–90.

Wilson, C. (2006). Learning Phonology With Substantive Bias: An Experimental and Computational Study of Velar Palatalization. *Cognitive Science*, *30*(5), 945–982.

Wonnacott, E. (2011). Balancing generalization and lexical conservatism: An artificial language study with child learners. *Journal of Memory and Language*, *65*(1), 1–14.

Wonnacott, E., Brown, H., & Nation, K. (2017). Skewing the evidence: The effect of input structure on child and adult learning of lexically based patterns in an artificial language. *Journal of Memory and Language*, *95*, 36–48.

Yu, A. C. L. (2004). Explaining final obstruent voicing in Lezgian: phonetics and history. *Language*, *80*(1), 73–97.