Phonological patterns in BEG-raising

Joseph A. Stanley
University of Georgia

Abstract Prevelar raising, or the raising of /ɛ/ before voiced velars, has been treated as a sound change that applies uniformly across relevant lexical items. Using self-reported responses from 434 participants, this paper shows that prevelar raising is more likely to occur if the velar is word-final (*beg*, *leg*) or intervocalic (*negative*, *legacy*) while raising is significantly less likely to occur when the velar is followed by a sonorant (*pregnant*, *regulate*), especially if that sonorant is a liquid (*negligent*, *segregate*). The set of words containing /ɛg/ is more heterogeneous than previously reported, suggesting that there are language-internal factors that must be controlled for when studying prevelar raising.

1 Introduction

1.1 Prevelar raising

In some varieties of North American English, certain front vowels are raised before voiced velar consonants—a phenomenon known as PREVELAR RAISING. This means that a speaker might say words like *back* and *tack* with the low vowel [æ] but pronounce the vowel in *bag* or *dragon* with a mid vowel [ɛ] or even [e]. Similarly, they may raise the vowels in words like *egg* or *legs* to [e] while keeping /ɛ/ in other environments lower. These two vowels, BAG and BEG respectively, approximate or may even merge with VAGUE, defined as /e/ before voiced velars as in *flagrant* or *bagel*. The related phenomenon of prevelar /ɪ/-raising in California (*e.g.* Cardoso et al. 2016) does not (yet) appear to be a robust or widespread change so it will not be discussed in this paper.

Recent scholarship has uncovered a high degree of variation in BAG-raising in the Northern United States and in the Pacific Northwest. In the first acoustic description of BAG-raising, Zeller (1997) finds variation in Milwaukee, with realizations ranging from [æ] to [e]. Raising is confirmed in other parts of Wisconsin, though not to the degree of merging with VAGUE or BEG (Bauer & Parker 2008). In addition to the North, BAG-raising is widespread in the Pacific Northwest, where it is robust across ages and ethnicities in Seattle and eastern Washington (Wassink

2016) and British Columbia (Swan 2015, 2016), often being associated with positive feelings toward the region (Swan 2018, Stanley 2018). It can also be found in Montana (Bar-El et al. 2017), other parts of Canada (Labov et al. 2006), and to a lesser degree as far south as San Francisco (Cardoso et al. 2016). This is not a uniquely American phenomenon: BAG-raising has been a feature of the English in Belfast, Ireland for over 150 years (Milroy & Milroy 1978, Milroy 1981, Patterson 1860).

However, compared to BAG-raising, there has been comparatively little research on BEG-raising. It is clear that the two are related because they nearly always co-occur. For example, BEG is completely merged with VAGUE both at the midpoint (Wassink et al. 2009, Wassink 2015) and the entire length of the vowels' trajectories (Freeman 2014, Riebold 2015) among BAG-raisers in Washington. However, some speakers in British Columbia appear to merge BAG and VAGUE while keeping BEG distinct (Mellesmoen 2018) and BEG-raising was found in Nevada without BAG-raising (Gunter et al. 2017). So, since some speakers have one and not the other, further investigation into BEG-raising is necessary to elucidate the status of prevelar raising in North American English.

1.2   Motivations for prevelar raising

The explanations for prevelar raising primarily have to do with language-internal factors. One primary argument is that the velar pinch, which is the raising of F2 and lowering of F3 in the transition formants into and from velar consonants, has also been cited as a possible motivation for prevelar raising even in words without tautosyllabic /g/ like *leggings* (Wassink & Riebold 2013). Baker et al. (2008) reinforce this by showing that the raised tongue body caused by the velar consonant correlates with the lowering of F1 and suggest that at some point this coarticulation was exaggerated. But velarity cannot be the only factor motivating prevelar raising because that would predict at least some raising before /k/ as well[1]. Because North American varieties of English do not have any indication of such pre-/k/ raising (Zeller 1997, Gunter et al. 2017), the velar pinch—if it even is a motivating factor—must be accompanied by some other cause.

---

[1]Actually, 19th century Belfast English raised /æ/ both before and after voiced and voiceless velars (Patterson 1860), so velarity can be the only motivation in some cases. But this does not seem to be the case in contemporary North American varieties of English.

Zeller (1997) argues that it is a combination of voicing, place of articulation, and nasality that lead to an articulatory motivation for the change. Voicing is a plausible contributing factor because vowels near voiced segments tend to have a lower F1 due to the slightly lower position of the larynx (Hillenbrand et al. 2001). Furthermore, since English vowels are longer before voiced segments, that extra length allows the velar pinch to be fully articulated. In the case of raising before /ŋ/, nasality is also a contributing factor because the opening of velum in anticipation of the nasal consonant causes a nasal formant in the vicinity of F1 to be interpreted as a lowered F1 (Olive et al. 1993). This combination of articulatory factors paints a plausible picture for the underlying cause for prevelar raising.

In addition to this articulatory motivation, the fact that there are so few words with BEG or VAGUE may be another reason for prevelar raising. While there are a handful of minimal pairs contrasting BAG and BEG (*drags-dregs*, *bag-beg*, *lag-leg*, etc.), there is just one known minimal pair contrasting VAGUE and BEG: *Prego* (the brand of pasta sauce) and the slang term *preggo* 'pregnant' (Stanley 2016). Figure 1 illustrates a play on words involving this minimal pair. However, one obscure pair is likely not enough to help retain the distinction and Zeller (1997) points out that

Figure 1   In an announcement of her pregnancy, this woman uses a jar of Prego sauce as a play on words with the term *preggo*, highlighting the minimal distinction between the two words. Used with permission.

the lack of minimal pairs between VAGUE and BAG allows a perceptual merging of the two sounds first before production. If the distinction between two sounds is not a robust tool for differentiating words, it can open the door towards merger (Wedel et al. 2013).

While language-internal factors are a compelling reason for prevelar raising, they cannot be the only motivations. One reason is simply the fact that other varieties of English do not have prevelar raising. Furthermore, Bauer & Parker (2008) point out that while there are very few minimal pairs between VAGUE and BAG, there are likewise very few that contrast VAGUE with /ig/ (*bagel* and *beagle* being one) but there is no evidence of them merging. Instead, Zeller (1997) argues that it is the lack of contrast in conjunction with the monophthongal quality of mid vowels in the northern Midwest (Thomas 1958) that opened the avenue for the loss of the phonemic contrast. It does seem to be the case that monophthongal mid vowels correlate with prevelar raising: both are found in contemporary Seattle speakers (Wassink 2015), and in southwest Washington prevelar raising and monophthongal /o/ were lost around the same time (Stanley 2017). If these language-internal effects are the primary motivation, Wassink & Riebold (2013) suggest that it is possible that this raising developed independently in Wisconsin and the Pacific Northwest.

1.3   Lexical patterns in prevelar raising

One language-internal factor that has yet to be explored in detail is that of lexical effects. To date, the majority of research on prevelar raising, and especially BEG-raising, has treated the entire lexical set as a uniform block of words and has not considered idiosyncratic patterns that individual members of the BEG class might have. Though Wassink & Riebold (2013) suggest anecdotally that *egg* may be raising nationwide as a case of a lexically isolated form of BEG-raising, this has not been tested empirically. It does appear that some words are raised more than others: Gunter et al. (2017) find that *egg* and *Peggy's* are raised much more often than *beg*, *keg*, *leg*, *leggings*, *peg*, and *pegs*. This suggests that there may be more heterogeneity this relatively small class of words than previously reported.

To date, few lexical effects have been found in BEG-raising possibly because only a small subset of BEG words are typically selected for analysis. There is merit in controlling for syllable structure or stress and keeping the number of words elicited from each participant small, but the effect is that most of what is known about BEG-raising has been primarily based on frequent monosyllables like *leg*,

peg, *keg*, and *egg*. What remains to be explored is the importance of word length and the possibility of longer words behaving differently from shorter words. When the entirely of a lexical class is small, getting a complete picture of the vowel is possible, and a robust analysis of the BEG vowel should include most of the words in the class rather than a small subset.

Anecdotally, these longer, less frequent words may in fact be different. I grew up in a suburb of St. Louis, Missouri, and while I have no other phonetic features characteristic of the North (including BAG-raising), I do have BEG-raising in my speech except in a few words like *integrity*, *regulate*, and *segregate*. It is noteworthy that all of these words have a sonorant following the /ɛg/ sequence. But a study that sampled my speech using only the common BEG words would miss these lexical exceptions entirely.

1.4   The current study

Since the BEG lexical set is about 75% the size of the BAG set[2] and because of the lack of BEG-specific studies, this paper focuses on BEG to test whether lexical effects are an important factor in prevelar raising. Specifically, and based on the patterns found in my own speech, I hypothesize that BEG-raising is less common in words that have a sonorant following /ɛg/. If the characteristics of my speech patterns extend beyond my idiolect, data would show that some other speakers with BEG-raising have similar exceptions. This contrasts with the null hypothesis which states that the presence of a following sonorant is not a significant predictor. Support for this null hypothesis would be found if the data shows no difference between /ɛg/ with following sonorants and /ɛg/ in other environments suggesting that any similarity in my list of exceptions is coincidental. If support for the main hypothesis is not found but differences within BEG words are still present, it may be that some other factor is involved or that there are purely idiosyncratic lexical differences. If there is no discernible difference between any BEG words, it may be the case that either they all are equally subject to raising or that each person has their own list of random exceptions and that no other language-internal factors are at play.

---

[2]The *Routledge Dictionary of Pronunciation for Current English* (Upton & Kretzschmar Jr. 2017) has 435 tokens with BEG and 589 with BAG. Similarly, the *CMU Pronouncing Dictionary* (Lenzo 2013) has 418 tokens with BEG and 560 with BAG.

2   Methods

To test this hypothesis, data from many lexical items with the BEG vowel in many phonological environments is required. In this section, I describe the methods for collecting this data using a categorization task distributed via an online survey.

2.1   Word selection

The words used in this study were selected by searching the 134,000-entry *CMU Pronouncing Dictionary*[3] (Lenzo 2013) for any word that contained the sequence 'EH1 G', which corresponds to a stressed /ɛ/ followed by a /g/. Of the 418 matches, 24 of the more frequent words were selected for inclusion loosely based on the following two rules:

i. Most proper names, which includes the majority of the tokens in the dictionary, were excluded. Such words that appeared in the dictionary that could have been included were *Clegg*, *Draeger*, *Egbert*, *Hegel*, *McGregor*, *Pegasus*, and *Peggy* but most of these would have been excluded based on frequency anyway. Additionally, they are possibly subject to idiolectal variation as well. For example, Gunter et al. (2017) found that *Peggy* was actually raised more than most other BEG words. To test other proper names, *Craig*, *Greg*, and *Lego* were selected because they appeared to be the most common.[4]

ii. Compounds and other derived forms were generally excluded, and singular nouns and unconjugated forms of verbs were preferred. Examples of such exclusions are *eggnog*, *eggplant*, *impregnate*, *irregular*, *kegs*, *megabyte*, *regulate*, and *segregated*. However, two derived forms were included: *leggings*, to contrast with *leg* and because of its inclusion in previous studies (Wassink & Riebold 2013), and *dregs* because the plural form is far more common than the singular (Davies 2008–).

---

[3]Freely available at http://www.speech.cs.cmu.edu/cgi-bin/cmudict.

[4]In retrospect, *Pegasus* should have been included because it is a somewhat well-known name. Also, to allow for maximal comparison with previous studies, *Peggy* should also have been included, but the fact that Gunter et al. (2017) included it in their study was only made known after the data collection for this project was complete.

Thus, conforming with these criteria, the following 24 BEG words were included in this survey: *beg*, *Craig*, *dregs*, *egg*, *Greg*, *integrity*, *interregnum*, *keg*, *leg*, *legacy*, *leggings*, *Lego*, *mega*, *negative*, *negligent*, *omega*, *oregano*, *peg*, *pregnant*, *regular*, *regulate*[5], *segment*, *segregate*, *segue*. The frequency of these words should be

[5]A reviewer pointed out the apparent redundancy in including *regular* and *regulate*. I have no reason for doing so other than my idiolect has a raised vowel in *regular* but not in *regulate* and I wanted to see if others had the same pattern.

word	COCA	log(COCA)
leg	65,689	4.82
regular	55,959	4.75
negative	44,354	4.65
egg	39,711	4.60
pregnant	30,922	4.49
Greg	24,886	4.40
segment	16,058	4.21
legacy	14,507	4.16
regulate	14,171	4.15
integrity	11,728	4.07
Craig	11,149	4.05
beg	5,325	3.73
peg	3,566	3.55
mega	3,417	3.53
segregate	2,965	3.47
omega	2,889	3.46
negligent	2,839	3.45
oregano	1,692	3.23
Lego	1,430	3.16
segue	1,021	3.01
keg	882	2.95
leggings	774	2.89
dregs	350	2.54
interregnum	108	2.03

Table 1   Frequency and log frequency of each lexical item.

considered because vowel shifts have been shown to affect more frequent words (Labov 1994). To gather data in the frequency of these lexical items, I searched for each word in COCA (Davies 2008–); because each lexeme has multiple word forms, I included the frequency of plurals and conjugated forms of verbs in the total frequency of that lexeme. For example, *beg* occurred 4,002 times and *begs* 1,323 times, so the total frequency of the lexeme *beg* for this paper was 5,325. Table 1 shows the frequency and the log (base 10) frequency for each of these 24 BEG words.

Even though the focus of this study was specifically on BEG-raising, additional words with other prevelar vowels were included in the survey, primarily to provide some variety to the questions and to serve as reference points for the other vowels. Six words with BAG (*bag*, *dragon*, *drags*, *flag*, *magnet*, and *snag*) and three VAGUE words (*vague*, *fragrance*, and *flagrant*) were selected for inclusion in the survey. The word *Prego* was also included, despite being a proper noun and its infrequency, because it is part of the only minimal pair with BEG. The words with BAG serve as a quick diagnostic for whether participants have BAG-raising and for the very small class of VAGUE words, it is assumed that speakers would use a higher [e] vowel, even if they do not have any form of prevelar raising.

2.2 Survey design

The online survey contained two main blocks of questions displayed as a single webpage. One block collected demographics about the participants and included fill-in-the-blank boxes asking for their ages and where they were from. The main block, which had the categorization task, listed each of the 34 aforementioned words with the orthographic letter of the targeted vowel underlined. For each word, participants were given the prompt, "How do you pronounce the highlighted sound?" and were then presented with these five choices in this order:

   i. The same vowel sound as 'BAKE'

  ii. The same vowel sound as 'DECK'

 iii. The same vowel sound as 'BACK'

 iv. Sometimes like 'BAKE" and sometimes like 'DECK'

  v. Something else (see comments section below)

The first three choices were used to refer to [e], [ɛ], and [æ] realizations. Option 4 allowed for an intermediate answer between [e] and [ɛ], indicating partial raising or oscillation between the two sounds. This is the option that people with some degree of BEG-raising—but not to the point of merging with VAGUE—would have presumably used[6]. Finally, the last option allowed for some other choice and encouraged elaboration in the comment box at the bottom of the survey, allowing for unforeseen realizations, idiosyncratic pronunciations, or other additional commentary. All 34 words in a different random order for each participant.

Option	Reference word	Approximate IPA	value
1	*bake*	[e]	1
4	(intermediate)	[e̞] or [ɛ̝]	0.5
2	*deck*	[ɛ]	0
3	*back*	[æ]	−1
5	(other)	NA	(excluded)

Table 2   Method for quantifying degree of raising based on participant responses.

To quantify degree of raising, the responses were converted into numbers as seen in Table 2. The higher the vowel, the higher the numeric value assigned to that response, with [e] being 1 and [æ] being −1. All 'other" responses (option 5) were not given a value and are excluded from analysis. With this method, if more speakers indicated a word to be raised, the word's average value would be closer to 1. If a BEG word was raised less often, it would have an average value closer to 0. In other words, a higher score indicated more raising.

The online survey was created and administered via Qualtrics and distributed over social media (primarily Facebook and Reddit). A total of 473 people took the survey, yielding 15,897 responses for the 34 words. Most participants were in their late 20s—reflecting my social network—and while most areas of the United States were included in the sample, there was a larger number of participants from Missouri, Utah, and Georgia (areas where I have lived), which are not reported to exhibit prevelar raising. Of these 473 people, 26 came from various countries outside North America; because of the small number from each country, they are

---

[6]This might have also been used by those with VAGUE-lowering (Bauer & Parker 2008, Freeman 2014, Gunter et al. 2017), but there was no indication of that here.

excluded from the main analysis of this paper. Therefore, the 434 self-reported North American participants are analyzed here. Due to the shallow age range, disproportionate geographic coverage, the failure to include gender or other demographic questions in the survey (an oversight in survey design), and the drawbacks due to a convenience sample, social factors will be left for future study when a more balanced dataset is available.

3   Results

The results from this survey show that there are phonological and lexical factors in self-reported BEG-raising. As mentioned above, words from three lexical classes were included in the survey, with the majority of them coming from the BEG lexical set. There was some slight variation in the VAGUE words, and the BAG words were more or less homogeneous in their responses. However, it is clear that there is systematic variation within the BEG words. Figure 2 shows all words, in order of how raised they were according to the quantification described in the previous section, and the number of people who chose each option is indicated by the color of the bars. The next three sections describe each of the vowels in the survey.

3.1   The VAGUE vowel

The top of Figure 2 shows the four words with VAGUE included in the survey and, as expected, it is clear that the vast majority of participants described these words as having the same vowel as in *bake*. For 96% of participants, the words *vague* and *fragrance* were reported as being said with [e]. The next word, *flagrant* was slightly lower since 37 people (8%) said they would use [æ] which may be a spelling pronunciation due to its infrequency and the orthographic <a>.[7] The one VAGUE word that stands out is *Prego*, which was [e] for roughly half the participants. This is discussed further below when taking into consideration other borrowings.

3.2   The BAG vowel

The bottom of Figure 2 shows the six words that belong in the BAG lexical class and, as expected, the vast majority indicated they would use the same vowel as

---

[7]In audio data collection unrelated to this project, I have included *flagrant* in word lists and some people actually do say *fl*[æ]*grant*.
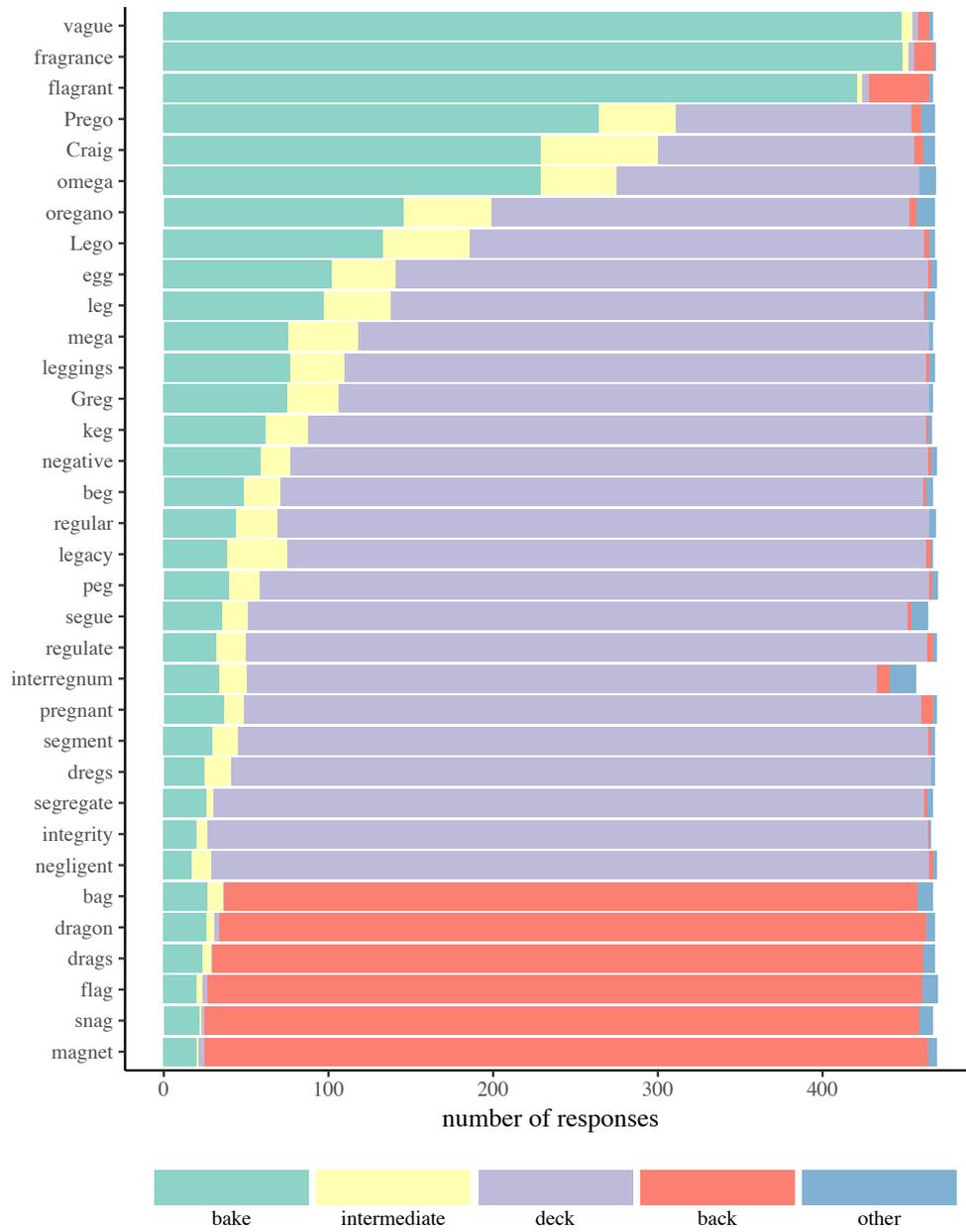
Figure 2   Number of responses for each option for each word, with words in order of their calculated degree of raising.

in *back*. However, BAG-raising was seen in about 23 people (5%) in this sample, and it appears it was done so approximately evenly across the six words. In other words, there does not seem to be any significant difference in how often *bag* and *dragon* were raised compared to *flag* and *magnet*, but this may be because of the lack of survey responses from BAG-raisers. The ranking of the six words does not seem to follow clear phonological rules other than *magnet* being the only one with a sonorant following the /g/. Since the current data do not provide any more insight, additional testing is required to see if manner of articulation of the following consonant is a significant factor in BAG-raising.

3.3   The BEG vowel

The most interesting aspect of Figure 2 is the large center portion which shows the BEG words. This section provides the strongest evidence for phonological effects of BEG-raising. It is immediately apparent that not all words raised equally often and that there is a continuum of how frequently raised each word tended to be. What stands out is that some words are raised far more than others, and a closer examination reveals that there are language-internal factors that correlate with this raising.

The BEG words that were raised by the most number of people are *Craig*, *omega*, *oregano*, and *Lego*. Additionally, the word *Prego*, which was considered a VAGUE word in this study, fits in better this group of words than with the rest of the VAGUE words. Each of these words are peculiar for their own reasons. The name *Craig* was the only word in this list with orthographic <ai> representing the prevelar vowel. *Omega* and *oregano* are relatively transparent borrowings from Greek and Italian, and *oregano* may be realized with stress on the penultimate syllable, which would skew the results. Finally, *Lego* and *Prego* are trademarks that originated from Danish and Italian words, respectively. About a third to half of the participants indicated the higher vowel in these five words, which were the most variable in this sample. For these people, it difficult to tell with the current data if these words and belong to the BEG lexical set and are being raised or if they have been rephonologized and are now part of the VAGUE class.

We can shed some light on this question by just analyzing the responses from the 26 people who indicated that they were from outside of North America. These participants strongly preferred [e] for *Craig* and *Prego*, but almost unanimously preferred [ɛ] for *Lego*, *Greg*, and the rest of the BEG lexical set. This suggests

that while there is little BEG-raising outside of North America (the sample did not include anyone from Belfast), *Craig* and *Prego* pattern with other VAGUE words for these people and that *Lego* and *Greg* belong to the BEG class of words.[8] With the exception of the British English pronunciation of *Craig* being listed with /eɪ/, both British and American pronunciations of *Greg* and *Lego* have /ɛ/ as their only variant. *Prego* was not listed in their dictionary.

It is impossible to say definitively what vowel was adopted when these words were borrowed into English, but because there are so few lexical items in the VAGUE lexical set, it is possible that these borrowings defaulted to the more frequent BEG. If that is the case, these borrowings undergo BEG-raising by the most number of people. Another possibility is that some people have VAGUE in these words and others have BEG. The question that remains is why *omega* and *Prego* were raised significantly more often than *oregano* or *Lego* (or even *mega*).

In the rest of the words with BEG, there is an almost perfect split that divides them: tokens where the /g/ is word-final or intervocalic were raised more than those that had a consonant following the /g/. The only exception was *regular*, which was raised slightly more than *legacy* and *peg*. This is compelling evidence for the proposed hypothesis that there is less BEG-raising in presonorant environments.

To formally test this split between presonorant and other environments, a generalized-linear mixed-effects model using the lme4 package (Bates et al. 2015) in R was fit to the data. The model included only the data for the BEG words, excluding the five borrowings and proper nouns discussed above, and only responses of [e] or [ɛ] to create a binary response. The indicated pronunciation was the dependent variable, and log frequency and the presence of a following sonorant were the independent variables, with participant and word as random intercepts. The model output (Appendix A, model 1) suggests that, as expected, lexical frequency is a significant factor and that more frequent lexical items were reported to be raised more often. However, after that is accounted for, this model shows that the presence of a following sonorant is significant and words with that phonological structure were raised often. This is evidence to reject the null hypothesis that manner of articulation of the following consonant has no effect on BEG-raising and supports the hypothesis that consonant after the /g/ can have an effect on the prevelar vowel.

While the finding that BEG words with following sonorants are raised less was expected given the pattern in my idiolect, one of the most surprising result from this

---

[8]This matches the pronunciations listed in *The Routledge Dictionary of Pronunciation for Current English* (Upton & Kretzschmar Jr. 2017).

data was that the manner of articulation of the following sonorant seems to have an effect as well. The difference was not drastic, but liquids (*negligent*, *integrity*, and *segregate*) had less raising than nasals (*segment*, *pregnant*, *interregnum*) which had less than glides (*regulate*, *segue*, *regular*). To test whether this pattern was statistically significant, a second generalized linear mixed-effects model was fit to the words containing a following sonorant, with the manner of articulation of the sonorant and log frequency as the fixed effects (see Appendix A, model 2). This time, lexical frequency was not statistically significant, and neither was the difference between glides and nasals, but liquids were raised significantly less often than either group. Because the lexical set containing BEG is relatively small, particularly those containing following sonorants, it may be coincidence that the BEG tokens with glides were raised more than those with nasals, but a larger dataset can confirm this. A future study that augments this list of lexemes with carefully constructed nonce words would reveal how robust and productive these patterns are as well.

Because there was only one word in the survey with a following obstruent (the /z/ in *dregs*), it is not feasible to determine how obstruents might fit into this hierarchy. Future studies should include derived forms such as *legs*, *pegs*, *kegs*, and *eggs* as well as the singular *dreg* to see if they pattern more with their base forms or if a following /-z/ negatively correlates with BEG-raising. However, a confounding variable here is that the /g/ and /z/ straddle a morpheme boundary in all of these words, and it is unknown whether morphological structure is a factor (though see Wassink 2016 for evidence that it is not an important predictor).

The rest of the BEG words are those containing a word-final or intervocalic /g/. Unlike the words with preconsonantal /g/, there is no obvious phonological pattern to explain why words like *peg* and *legacy* would be raised relatively less often compared to *leg* and *egg*. The words with intervocalic /g/ are interspersed among those with word-final /g/, so it is not simply a factor of the /g/ being tautosyllabic. A third generalized linear mixed-effects model was fit to these words (excluding the borrowings discussed above and the presonorant tokens) with the environment (word-final or intervocalic /g/) and log frequency as the fixed effects (Appendix A, model 3) and found that neither lexical frequency nor the environment were statistically significant factors. Unlike the presonorants, there is a fairly wide range of how often words within this group were raised, and it is clear that some words are raised by more people than others.

4   Discussion

The previous section has shown that prevelar raising does not happen uniformly across the lexicon but that some words, particularly those with a following sonorant, are less likely to be raised than others. With this in mind, we can reject the null hypothesis that all words with BEG are equally likely to be raised. It may still be the case that each person has their own list of exceptions, but this list may be grounded in phonological properties of the word rather than a random subset of words containing BEG.

An interesting point of comparison that should be made is that Gunter et al. (2017) actually included one presonorant BEG word, *regular*, in their study. In fact, they find that *regular* was nearly always raised so it was removed from some of their analyses and forms part of their defense for lexical motivations for this merger. *Regular* is the second most frequent BEG word in this study, and the most frequent in the presonorant environment, so it comes as no surprise that it is often raised. An anonymous reviewer points out that a form of long-distance assimilation or harmony may be in effect for this word and that the palatal glide itself may be the cause of this raised vowel. In contrast, while *regular* was the most raised presonorant token in this sample, it was not found to be categorically raised to [e] as it was in found in Nevada. It may be that BEG-raising is spreading across the lexicon differently in Nevada than it is in other parts of the country. This discrepancy only increases the need for the inclusion of these words in prevelar studies.

It is impossible to say with this data whether BEG is actually merging with VAGUE, either in production or perception. However, if such a merger is happening, it is important to consider what type of merger it may be (Labov 1994). It does not appear to be happening via merger by approximation (Trudgill & Foxcroft 1978) where the BEG and VAGUE gradually approach each other in phonetic space. If this were happening, not only could the number of intermediate responses be higher, but the entire lexical class would raise *en bloc* rather than being systematically raised by phonetic conditions. It likewise does not appear to be a merger by expansion (Herold 1990), where the phonetic space of the merged vowel occupies the same space as the union of the two historic classes. Evidence for this kind of merger would come in the form of the VAGUE vowel being lowered more frequently, particularly by those with some degree of BEG-raising, which was not found in this sample. A possible candidate is that this is a merger by transfer (Trudgill &

Foxcroft 1978), where the lexical items from one lexical set abruptly defect to and are rephonologized as being part of a different vowel. However, the issue here is that this type of merger is presumably blind to other phonetic factors and that the words would shift haphazardly and one at a time. This would not account for the phonological conditioning found in this data.

Instead, if the BEG is indeed merging with VAGUE, it appears to be what Dinkin (2016) calls a merger by 'phonological transfer.' With this type of merger, 'rather than individual lexical items being transferred one at a time from one phonemic class to the other, entire sets of words, determined by the phonological environment of the merging phoneme, are transferred simultaneously' (Dinkin 2016: 164). Thus, instead of a change percolating across the lexical set one at a time, a merger by phonological transfer posits that a change traverses the lexicon stepwise, with entire phonological environments being affected at once. This type of merger accounts for the patterns in BEG-raising in a way that the others do not. This pattern is similar to the merging of BAG and VAGUE in some speakers in British Columbia, which has also been described as a merger by phonological transfer (Mellesmoen 2018). Under the hypothesis of a merger by transfer, BEG-raising affected words where the /g/ is intervocalic or word-final first. For some speakers, but not all, it has spread to presonorant environments as well, with some indication that specific natural classes within the sonorants are affected before others. What makes the pattern of BEG-raising different from the phonological transfer described by Dinkin (2016) is that the more general phonological environment seems to have undergone the change, leaving a more restricted environment behind. Even if BEG is not completely merging with VAGUE, the properties of a merger by phonological transfer are still applicable to BEG-raising as a description of sound change more generally rather than as a theory of merger.

An important implication of the phonological effect in BEG-raising is in the construction of word lists and selection of tokens to include for analysis. Previous studies of BEG-raising have carefully controlled for syllable and morphological structure and included a variety of words with BEG preceding tautosyllabic /g/ (*peg*, *beg*, *egg*, *leg*) and intervocalic /g/ (*negative*, *leggings*) (Freeman 2014, Gunter et al. 2017). However, given the finding that these are the very words that are more likely to be raised, excluding other words may overestimate the degree of BEG-raising and may not reveal the full pattern (*i.e.* whether BEG is raised before presonorant /g/ as well). This study suggests that presonorant /g/ should also be added to the list of environments that a study might control for.

One unanswered question is what the relationship must be between the /g/ and the following consonant. In the current study, they were all within the same morpheme with the exception of *dregs*. Because plurals were generally avoided in this study, there is no data on *leg* versus *legs* or *egg* versus *eggs*. Additionally, compounds such as *eggnog*, *eggplant*, *eggshell*, *legwork*, *legroom*, or other words like *legless* or *Egbert* were not included. These may shed additional light on how morphological boundaries affect BEG-raising and should also be considered for analysis in prevelar studies in order to get a more complete picture of the phenomenon.

5   Conclusion

This paper has provided evidence for phonological effects in BEG-raising in North American English speakers using data collected from an online categorization task. Specifically, when /ɛg/ is followed by a sonorant, the vowel is less likely to raise than in words with an intervocalic or word-final /g/. Furthermore, the manner of articulation of the following sonorant may also be an important factor with words with following liquids being raised less frequently than glides (within the presonorant BEG tokens). There was also evidence that transparent borrowings and proper nouns were much more likely to be raised. This is the first study to show a long-distance phonological effect in BEG-raising and possibly prevelar raising generally.

Because of the nature of the data collection instrument, there are some obvious limitations to this study. Self-reported data on pronunciation is not completely reliable and further acoustic, articulatory, and perceptual data are necessary to confirm the phonological patterns found here. Because the number of tokens with BEG is so small, finding enough tokens for the various sonorants may not be feasible without resorting to nonce words. Also, since the regional distribution of BEG-raising is primarily limited to the North and the Pacific Northwest, additional work is needed to uncover regional and sociolinguistic patterns associated with BEG-raising.

This study has shown that the set of words containing BEG is less homogeneous than has been reported in the past. These findings have implications for how prevelars should be analyzed in future studies and open the door to additional work. The next step is to uncover the motivations for why a following sonorant would inhibit raising and to see if the patterns for BEG-raising extend to other prevelar

vowels as well. BAG-raising is currently a fruitful area of research, but with the addition of BEG-raising as a distinct linguistic phenomenon, the study of prevelars will continue to shed more light on conditioned sound changes in North American English.

Appendix: Model summaries

This appendix contains the model summaries for three generalized linear mixed models fit by maximum likelihood (Laplace Approximation) and their interpretation. For each of these models, only BEG words are included except for the five borrowings and proper nouns were excluded (*Prego*, *omega*, *Craig*, *oregano*, and *Lego*). I also excluded any response that was not either *bake* or *deck* in order to make the response variable binary. The response variable was *bake* vs. *deck*, and *deck* was coded as the reference value because the lower vowel appears to be the underlying and historic form and so that the positive coefficients in the output indicate raising.

Model 1: Presonorant verses other environments

The data for this model included all 8,325 responses from BEG words. The fixed effects were log-transformed lexical frequency and whether the /g/ was followed by a sonorant, with word and participant as random intercepts. The number of syllables as an ordinal variable was initially included in the model but it resulted in an overfit model (it was only significant as a cubic effect) and was excluded.

	Estimate	Std. Error	z-Value	p-Value	
Intercept	−9.189	1.109	−8.288	< 0.001	∗∗∗
log_frequency	0.438	0.205	2.141	0.032	∗
is_presonorant (=true)	−1.905	0.328	−5.808	< 0.001	∗∗∗

Random Effects	Variance	Std. Dev.
Word	0.404	0.636
Participant	39.029	6.247

Table 3 Presonorant verses other environments. Formula: response ~log_frequency + is_presonorant + (1|word) + (1|participant

This table shows the model output for a generalized linear mixed-effects model fit by maximum likelihood. The negative coefficient for the intercept reflects the fact that most of the BEG words were more often reported as [ɛ] instead of [e]. The positive coefficient for the log frequency suggests that as the frequency increases, the likelihood of that word being raised increases. The negative coefficient for when is_presonorant = true suggests that presonorant tokens of BEG were less likely to be raised.

Model 2: Place of articulation for following sonorant

The data for this model included 1,218 tokens of BEG where the /g/ was followed by a sonorant. Not all participants had observations of bake in this environment (271 categorically indicated *deck* before sonorants), so this model includes data only from those with some variation (n = 179). This had no effect on the overall results, though the coefficients are slightly smaller. The fixed effects were log-transformed lexical frequency the manner of articulation of the following sonorant: nasals (*interregnum*, *segment*, *pregnant*), liquids (*integrity*, *segregate*, *negligent*), and liquids (*regular*, *regulate*, *segue*). The reference level was set to be the nasals since they appeared in the middle in Figure 2. The number of syllables was initially included in this model as a fixed, ordinal variable, but the model failed to converge.

	Estimate	Std. Error	z-Value	p-Value	
Intercept	-2.202	0.521	-4.230	< 0.001	$***$
log_frequency	-0.045	0.132	-0.342	0.732	
manner (=glide)	0.254	0.264	0.964	0.335	
manner (=liquid)	-1.058	0.276	-3.840	< 0.001	$***$

Random Effects	Variance	Std. Dev.
Word	0.003	0.050
Participant	5.240	2.289

Table 4   Place of articulation for following sonorant. Formula: response ~log_frequency + manner + (1|word) + (1|participant)

This table shows the model output for a generalized linear mixed-effects model fit by maximum likelihood. The negative coefficient for the intercept reflects the

fact that most of the BEG words were more often reported as [ɛ] instead of [e]. The log frequency was not a significant predictor in this model. The coefficient for the glides is positive but small, suggesting they are raised more, but the large *p*-value indicates that this difference is not statistically significant. However, the coefficients for the liquids was negative and statistically significant, supporting the finding that liquids were raised even less than nasals (and glides).

Model 3: Word-final verses intervocalic /g/

The data for this model included 1,979 tokens of BEG where the /g/ was either intervocalic or word-final. As with Model 2, only participants with non-categorical responses in this environment are included ($n = 236$). The fixed effects were log frequency and the environment of the following /g/: word-final (*beg*, *egg*, *Greg*, *keg*, *leg*, *peg*) or intervocalic (*legacy*, *leggings*, *mega*, *negative*). The reference level was set to word-final tokens. The number of syllables was not included in this model because it was strongly correlated with other variables: all word-final tokens had just one syllable and intervocalic by necessity had more than one, and both two-syllable words were less frequent than both three-syllable words.

	Estimate	Std. Error	*z*-Value	*p*-Value
Intercept	−2.967	1.901	−1.561	0.118
log_frequency	0.409	0.461	0.887	0.375
environment (=intervocalic)	−0.243	0.627	−0.388	0.698

Random Effects	Variance	Std. Dev.
Word	0.877	0.936
Participant	7.995	2.828

Table 5   Word-final verses intervocalic /g/. Formula: response ~log_frequency + environment + (1|word) + (1|participant)

This table shows the model output for a generalized linear mixed-effects model fit by maximum likelihood. The large *p*-values for both predictor variables indicates that neither log frequency nor environment were significant factors in predicting reported BEG-raising.

References

Baker, Adam, Jeff Mielke & Diana Archangeli. 2008. More velar than /g/: Consonant coarticulation as a cause of diphthongization. In Charles B. Chang & Hannah J. Haynie (eds.), *Proceedings of the 26th West Coast Conference on Formal Linguistics*, 60–68. Somerville, MA: Cascadilla Proceedings Project.

Bar-El, Leora, Laura Felton Rosulek & Lisa Sprowls. 2017. Montana English and its place in the West. In Valerie Fridland, Alicia Beckford Wassink, Tyler Kendall & Besty E. Evans (eds.), *Speech in the Western States, Volume 2: Inland West* (Publication of the American Dialect Society. 102), 107–138. Durham, NC: Duke University Press. https://doi.org/10.1215/00031283-4295299.

Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. https://doi.org/10.18637/jss.v067.i01.

Bauer, Matt & Frank Parker. 2008. /æ/-raising in Wisconsin English. *American Speech* 83(4). 403–431. https://doi.org/10.1215/00031283-2008-029.

Cardoso, Amanda, Lauren Hall-Lew, Yova Kementchedjhieva & Ruaridh Purse. 2016. Between California and the Pacific Northwest: The front lax vowels in San Francisco English. In Valerie Fridland, Tyler Kendall, Betsy E. Evans & Alicia Beckford Wassink (eds.), *Speech in the Western States: Volume 1: The Coastal States* (Publication of the American Dialect Society 101), 33–54. Durham, NC: Duke University Press. https://doi.org/10.1215/00031283-3772890.

Davies, Mark. 2008–. The Corpus of Contemporary American English (COCA): 560 million words, 1990-present. Available online at https://corpus.byu.edu/coca/.

Dinkin, Aaron J. 2016. Phonological transfer as a forerunner of merger in Upstate New York. *Journal of English Linguistics* 44(2). 162–188. https://doi.org/10.1177/0075424216634795.

Freeman, Valerie. 2014. Bag, Beg, Bagel: Prevelar raising and merger in Pacific Northwest English. *University of Washington Working Papers in Linguistics (UWWPL)* 32.

Gunter, Kaylynn, Ian Clayton & Valerie Fridland. 2017. Pre-velar raising and categorization in Nevada English. Third Annual Meeting of the NorthWest Phonetics & Phonology Conference, Vancouver, BC.

Herold, Ruth. 1990. *Mechanisms of merger: The implementation and distribution of the low back merger in Eastern Pennsylvania*. Philadelphia: University of Pennsylvania dissertation.

Hillenbrand, James M., Michael J. Clark & Terrance M. Nearey. 2001. Effects of consonant environment on vowel formant patterns. *The Journal of the Acoustical Society of America* 109. 748–763. https://doi.org/10.1121/1.1337959.

Labov, William. 1994. *Principles of linguistic change. Vol. 1: Internal features*. Oxford: Wiley-Blackwell.

Labov, William, Sharon Ash & Charles Boberg. 2006. *The atlas of North American English: Phonetics, phonology and sound change*. Walter de Gruyter.

Lenzo, Kevin. 2013. The CMU pronouncing dictionary.

Mellesmoen, Gloria. 2018. A remedial path to merger: Merger by phonological transfer in British Columbia English. *Toronto Working Papers in Linguistics* 40.

Milroy, James. 1981. *Regional accents of English: Belfast*. Belfast: Blackstaff Press Ltd.

Milroy, James & Lesley Milroy. 1978. Belfast: Change and variation in an urban vernacular. In Peter Trudgill (ed.), *Sociolinguistic patterns in British English*, 19–36. London: Edward Arnold.

Olive, Joseph P., Alice Greenwood & John Coleman. 1993. *Acoustics of American English speech: A dynamic approach*. Murray Hill, NJ: Springer-Verlag.

Patterson, David. 1860. *The provincialisms of Belfast and the surrounding districts*. Belfast.

Riebold, John Matthew. 2015. *The social distribution of a regional change: /æg, ɛg, eg/ in Washington State*. Seattle: University of Washington dissertation.

Stanley, Joseph A. 2016. Pacific Northwest English: Historical overview and current directions. *The University of Georgia Working Papers in Linguistics* 3.

Stanley, Joseph A. 2017. Changes in the timber industry as a catalyst for linguistic change. Poster presentation at the 46th New Ways of Analyzing Variation conference (NWAV46), Madison, WI.

Stanley, Joseph A. 2018. Changes in the timber industry as a catastrophic event: BAG-raising in Cowlitz County, Washington. *University of Pennsylvania Working Papers in Linguistics* 24(2). 403–431.

Swan, Julia Thomas. 2015. Raising or retraction? The story of /æ/ in Seattle and Vancouver, B.C. Presentation at the New Ways of Analyzing Variation 44, Toronto.

Swan, Julia Thomas. 2016. Canadian English in the Pacific Northwest: A phonetic comparison of Vancouver, BC and Seattle, WA. *Proceedings of the 2016 Annual Conference of the Canadian Linguistic Association* .

Swan, Julia Thomas. 2018. Toward a social meaning of BAG-raising: Sociohistorical factors and ideological stance among Seattle and Vancouver talkers. Presentation at the Cascadia Workshop in Sociolinguistics (CWSL3), Portland, Oregon.

Thomas, Charles Kenneth. 1958. *An introduction to the phonetics of American English*. New York 2nd edn.

Trudgill, Peter & Tina Foxcroft. 1978. On the sociolinguistics of vocalic mergers: Transfer and approximation in East Anglia. In Peter Trudgill (ed.), *Sociolinguistic patterns in British English*, 69–79. London: Edward Arnold.

Upton, Clive & William A. Kretzschmar Jr. 2017. *The Routledge dictionary of pronunciation for Current English*. London; New York: Routledge 2nd edn.

Wassink, Alicia Beckford. 2015. Sociolinguistic patterns in Seattle English. *Language Variation and Change* 27(1). 31–58. https://doi.org/10.1017/S0954394514000234.

Wassink, Alicia Beckford. 2016. The vowels of Washington State. In Valerie Fridland, Tyler Kendall, Betsy E. Evans & Alicia Beckford Wassink (eds.), *Speech in the Western States: Volume 1: The Coastal States* (Publication of the American Dialect Society 101), 77–105. Durham, NC: Duke University Press. https://doi.org/10.1215/00031283-3772912.

Wassink, Alicia Beckford & John Matthew Riebold. 2013. Individual variation and linguistic innovation in the American Pacific Northwest. Chicago Linguistic Society (CLS 49) Workshop on Sound Change Actuation, Chicago.

Wassink, Alicia Beckford, Robert Squizzero, Mike Scanlon, Rachel Schirra & Jeff Conn. 2009. Effects of style and gender on fronting and raising of /æ/, /eː/ and /ɛ/ before /ɡ/ in Seattle English. New Ways of Analyzing Variation (NWAV) 38, Ottawa.

Wedel, Andrew, Abby Kaplan & Scott Jackson. 2013. High functional load inhibits phonological contrast loss: A corpus study. *Cognition* 128(2). 179–186. https://doi.org/10.1016/j.cognition.2013.03.002.

Zeller, Christine. 1997. The investigation of a sound change in progress: /æ/ to /e/ in Midwestern American English. *American Speech* 25(2). 142–155. https://doi.org/10.1177/007542429702500207.