1   REGIONAL PATTERNS IN PREVELAR RAISING

2   Joseph A. Stanley

3   *Brigham Young University*

4

5

6

7   **Abstract:**   Prevelar raising is the raising of TRAP and DRESS vowels before voiced velars. While BAG-

8   and BEG-raising have been described in Canada, the Upper Midwest, and the Pacific Northwest, an in-depth

9   investigation of their distribution across North America is lacking, especially for BEG. Using an online

10   survey distributed to over 5,000 participants via Reddit and ordinary kriging for spatial interpolation, this

11   study finds that prevelar raising is more widespread than previously reported. BAG-raising is found in much

12   of the North and the Upper Midwest. BEG-raising is far more variable and is common across much of the

13   Midlands and the West, with concentrated pockets in the Northern Great Plains and various other regions.

14   This data suggests that the two can occur independently, with areas like the upper Midwest exhibiting BAG-

15   raising alone, and the Midlands and the West reporting BEG-raising alone. These findings suggest that

16   additional research on prevelar raising and other infrequent phonological variables is required to uncover

17   their regional distribution and social meaning.

18

1   Introduction

19

20   Prevelar raising is the raising of canonical TRAP and DRESS vowels before voiced velars.[i] Some speakers

21   of North American English pronounce the vowels in words like *bag*, *flag*, and *dragon* with a raised nucleus

22   ([ɛː] or [eː]) or off-glide ([æᶦ]). Similarly, some speakers pronounce the prevelar vowel in words like *beg*,

23   *legs*, or *integrity* as [ɛᶦ], [eː], or [eᶦ]. For consistency with previous research, the terms BAG and BEG will be

24   used, respectively, to refer to TRAP and DRESS in these environments. The process of raising these vowels

25   will be referred to as BAG-raising and BEG-raising, which both fall under the umbrella of prevelar raising.

26   While the amount of research on prevelar raising is growing, there are some limitations in the scope of

27   much of this research. For example, prevelar raising has only been described in detail in some parts of North

28   America, like Canada, the Upper Midwest, and the Pacific Northwest so the degree of raising outside of

29   these areas is largely unknown. Furthermore, most studies are based on a relatively small sample of words,

30   making it difficult to extrapolate to other words in the lexical set. This study overcomes these obstacles by

31   presenting reported raising on dozens of BAG and BEG words, gathered via online surveys from 5,269 people

32   from most of English-speaking North America. While there is some unreliability inherent in self-reports

33   such as these, the findings solidify the established dialect boundary for BAG-raising and suggest that BEG-

34   raising is far more widespread than has been documented. Furthermore, the data reveals areas where BEG-

35   raising occurs without BAG-raising and vice versa. The inclusion of many more words in this online survey

36   was justified as well since the amount of raising, particularly for BEG, is determined by which types of

37   words get studied.

2   Prevelar raising in North American English

38

39   Prevelar raising is an ongoing phenomenon in North American English, though there are scattered reports

40   of the phenomenon in much earlier studies. Perhaps the first account of prevelar raising is Patterson's (1860)

41   list of words that are "incorrectly" pronounced in Belfast, Ireland. The TRAP vowel is described a being

42   pronounced with "the short sounds of *e*", which refers to the lax [ɛ]: "this error is almost exclusively

43   confined to those words in which *a* is preceded by *e* or *g*, or followed by the sound of *k*, hard *g*, or *ng*"

44   (Patterson 1860: 15). Thus, TRAP among that community was raised to some higher vowel before and after

45   both voiced and voiceless velar sounds—a more encompassing environment than today's prevelar raising.

46   In North American English, perhaps the first description of prevelar raising is based on the *Thomas*

47   *Collection* (Thomas 1958). This collection, which is based on over 15,000 recordings gathered in the 1930s

48   and 1940s from the majority of the United States, establishes that prevelar raising has been established for

49   several generations.

50

51   Minor variations of other vowels also occur before [g] and [ŋ], though without any clear regional

52   pattern. Thus [æ], which is normally lax, may become tense before [g] or [ŋ], and the tension may

53   induce a raising of the tongue, as in [beg] for *bag*, [eŋgri] for *angry*, and [eŋkɚ] for *anchor*. The same

54   increase in tension may be noted in the change of [ɛ] to [e] in such words as *egg*, *leg*, and *beg*. (Thomas

55   1958: 204)

56

57   Thomas' early description hints at the pan-regional nature of the pattern and the link between BAG- and

58   BEG-raising. Upglides in words like *beg* and *bag* are also found in the Pacific Northwest (Reed 1961),

59   Indiana (Carmony 1970), and the South (Wells 1982: 531), confirming the widespread distribution of

60   prevelar raising.

61   Today, additional research has uncovered phonetic, regional, and social patterns in BAG-raising. In the

62   first dedicated study on BAG-raising, Zeller (1997) describes variation in Milwaukee, Wisconsin. Drawing

63   from a sample of 10 speakers, including her own family members, Zeller shows that some speakers (like

64   her father) have an underlying /æ/ in BAG, other people (like her mother) realize BAG with an intermediate

65   vowel quality between [e] and [ɛ], and a third group (including her brother) have a vowel indistinguishable

66   from [ɛ]. Bauer & Parker (2008) use experimental articulatory data to show that while BAG was raised, it

67   was demonstrably different from [æ], [ɛ], and [e], concluding that BAG was not merged with any other

68   vowel (see Stanley 2020 for a similar finding in Washington). Whether mergers are occurring when BAG

69   or BEG raise remains an open question in some regions, but it will not be a focus of this paper.

3

70   BAG-raising has been described in other northern areas of North America. In addition to Wisconsin

71   (Zeller 1997; Bauer & Parker 2008; Purnell 2008; Benson, Fox & Balkman 2011), it is found in Montana

72   (Bar-El, Rosulek & Sprowls 2017) and Alberta (Jones 2015; Rosen & Skriver 2015). In fact, among

73   Canadians, it is strongest in the Prairies (i.e. Alberta, Saskatchewan, and Manitoba; Boberg 2008: 146),

74   though it is reported in Vancouver as well (Swan 2016; Mellesmoen 2018). However, BAG-raising is most

75   thoroughly described in the Pacific Northwest (Wassink et al. 2009; Wassink 2015); it is found among

76   various age groups in Oregon (McLarty, Kendall & Farrington 2016; Becker et al. 2016) and ethnic groups

77   in Seattle (Riebold 2015; Wassink 2016) and there are detailed accounts of its social meaning in Seattle,

78   Vancouver, and Cowlitz County (Swan 2020; Stanley 2018). Generalizing the region where BAG-raising is

79   found, the *Atlas of North American English* (Labov, Ash & Boberg 2006) shows that speakers from a large

80   area—all of Canada west of Ottawa and the Midwestern and Western US states that border Canada—had

81   some indication of BAG-raising.

82   In contrast to this large amount of research on BAG-raising, far less is known about BEG-raising, the

83   bulk of which comes from the Pacific Northwest and surrounding areas. In Seattle, it is found in speakers

84   of all ages (Wassink 2015) who realize BEG with considerable phonetic overlap with VAGUE (that is, /eg/;

85   see below; Freeman 2014; Riebold 2015). In Portland, BEG-raising is found primarily in older speakers

86   (Becker et al. 2016). In Reno, BEG-raising is found in both White and Native American speakers (Gunter,

87   Clayton & Fridland 2017; Clayton & Fridland 2020) and may be lexically motivated since some words

88   were raised more consistently than others. Other than brief references to its presence in the South (Wells

89   1982: 531; Yavaş 2011: 83), BEG-raising has, to my knowledge, not been reported outside of the Northwest

90   and Nevada, possibly since few researchers have investigated BEG-raising in other areas. Therefore, that

91   BEG-raising is considered a Pacific Northwest phenomenon only reflects a sampling bias in the literature

92   since it has yet to be confirmed that BEG-raising does *not* occur elsewhere in North America.

93   Related to BAG- and BEG-raising is the very small set of words with /eg/ including *vague*, *fragrance*,

94   and *pagan*. In this paper, these words and this vowel will be referred to as VAGUE. Most North American

95   English speakers use /e/ in these lexical items, even if they do not exhibit any BAG- or BEG-raising. VAGUE

4

96 is relevant here because BEG and BAG appear to be merging with VAGUE in some communities (Freeman

97 2014; Wassink 2016). In fact, for speakers with a high degree of prevelar raising, VAGUE *lowers* so that

98 BEG and VAGUE merge somewhere between canonical FACE and DRESS (Freeman 2014; Riebold 2015).

99 Because there are so few words that contain VAGUE, it is an especially difficult lexical set to study, but it

100 provides a useful reference point when quantifying the degree of raising for BEG and BAG.

101 From the limited studies on these phenomena, we see that both BAG-raising and BEG-raising seem to

102 be related because they almost always occur together. However, as mentioned previously, all studies on

103 BEG-raising have been based in areas where BAG-raising is known to occur, so it may be the case that their

104 co-occurrence is limited to the Pacific Northwest. Stanley (2019) shows that BEG- and BAG-raising occur

105 in different phonological, morphological, and lexical environments, suggesting some independence.

106 Furthermore, there are documented cases of one without the other: some speakers in Nevada raise BEG

107 without raising BAG (Gunter, Clayton & Fridland 2017; Clayton & Fridland 2020) and some speakers in

108 British Columbia raise BAG without raising BEG (Mellesmoen 2018).[ii] This small but growing body of

109 research suggests some independence between BEG and BAG and that there are areas outside of Canada and

110 the greater Pacific Northwest region that may have some prevelar raising.

111 Turning now to the topic of methodology, it is important to consider how prevelar raising has been

112 studied in previous research. A linguistic variable is easiest to study when it is frequent, or rather, when it

113 "occurs so often in the course of undirected natural conversation that its behavior can be charted from

114 unstructured contexts and brief interviews" (Labov 1963: 179). Neither BAG nor BEG are frequent linguistic

115 variables. Cardoso et al. (2016) were forced to exclude both vowels from analysis because their corpus of

116 22 interviews contained just 15 tokens of each vowel. Wassink & Hargus (2020) heard just six tokens of

117 /ɛg/ and /ɛŋ/ (as in *strength*) in nine sociolinguistic interviews. Riebold (2015) and Swan (2020) likewise

118 heard few tokens in interviews and resorted to using supplemental wordlists to gather enough prevelar

119 tokens for analysis. Jones (2015) used pictures and the elicitation of nonce words to collect the vowel in

120 environments that, due to historical accidents, are not found in the English lexicon, like /vɛg/, /dɛg/, and

121 /pæg/. For such infrequent phonological variables, these types of targeted approaches illustrate the need for

122 new ways to collect enough data for a robust analysis.[iii]

123   Because of their infrequency then, most researchers have only been able to study BEG and BAG through

124 elicitations—that is, wordlists and reading passages—rather than via naturally occurring data. In addition

125 to inevitable, confounding stylistic effects that underlie these findings, most of what is known about

126 prevelar raising is based on relatively few word types. Specifically, BEG is most often analyzed by eliciting

127 tokens of *beg*, *egg*, *keg*, *leg*, *Meg*, *peg*, *leggings*, *regular*, *negative*, and *Peggy's*, with much overlap between

128 studies (Bauer & Parker 2008; Clayton & Fridland 2020; Freeman 2014; Gunter, Clayton & Fridland 2017;

129 Riebold 2015; Reed 1952; Swan 2016; Wassink 2016). For BAG, it is more diffuse, but most studies include

130 a subset of no more than eight of the following *bag*, *Bagdad*, *brag*, *crag*, *drag(ging)*, *dragon*, *flag*, *gag*,

131 *hag*, *haggle*, *lag*, *magnet*, *nag*, *pragmatic*, *sag*, *stag*, *tag*, and *(zig)zag* (Bauer & Parker 2008; Clayton &

132 Fridland 2020; Freeman 2014; Mielke, Carignan & Thomas 2017; Riebold 2015; Rosen & Skriver 2015;

133 Swan 2016; Wassink 2016). The notable exceptions are Jones (2015) and Stanley (2019) who expand their

134 lists to include a great number of additional words. This limited set of words gives a restricted view of

135 prevelar raising and may overreport or even underreport the amount of actual raising a person may have.

136   The purpose of this paper therefore is to address two questions. First, how widespread is prevelar raising

137 geographically? There is some indication that it is a pan-regional phenomenon, but a lack of focused

138 research and a shortage of data has prevented any confirmation of Thomas' (1958) early findings. This

139 study answers this question by sampling from all parts of English-speaking North America and confirms

140 that BEG-raising is far more widespread than previously reported. This study also addresses the related

141 question of independence between BEG- and BAG-raising and shows that there are regions that have one

142 without the other.

143   Second, how widespread is prevelar raising across the lexicon? Because BEG and BAG are such small

144 lexical sets, naturally occurring data usually does not warrant enough tokens for a robust analysis. However,

145 elicitations in previous research have only included a handful of all possible prevelar words. This study

146 tackles this issue head-on by gathering data from dozens of prevelar tokens, including many words that are

147   not typically included in wordlists, and shows that an expanded wordlist produces different results than a

148   restricted subset.

149   **3   Methods**

150   As we are now well into the 21st Century, researchers are exploring new ways to reach wider audiences

151   in dialectology research. Kim et al. (2019) analyzed the speech of 626 New Englanders collected over a

152   few months via Amazon Mechanical Turk and present results comparable to those found using data

153   collected using more traditional techniques. With the preponderance of smartphones today, mobile apps

154   have been created to crowdsource dialect data on Swiss German (Leemann et al. 2015) and British English

155   (Leemann, Kolly & Britain 2018), showing that a large amount of data can be quickly collected when taking

156   advantage of current technology. These studies show that despite their obvious shortcomings (a lack of

157   face-to-face interaction, for example) online, asynchronous, crowdsourced methods for data collection can

158   still produce useful results in dialectology research. This section describes the methods used to collect, and

159   subsequently analyze, a large amount of data on prevelar words. All data processing, including

160   summarization for the tables, was done in R using various packages and functions within the tidyverse

161   (Wickham 2017).

162   **3.1 Word selection**

163   The first task in data collection was to establish a comprehensive list of all prevelar words. For this

164   purpose, two dictionaries of English transcriptions were consulted. The first was the *CMU Pronouncing*

165   *Dictionary*[iv] (Lenzo 2013), an open source dictionary containing over 134,000 searchable transcriptions in

166   a machine-readable format. Using the pattern "AE1 G", "EH1 G", and "EY1 G" as search parameters, there

167   were 560 entries for BAG, 418 for BEG, and 223 for VAGUE, respectively. The second dictionary was the

168   *Routledge Dictionary of Pronunciation for Current English* (Upton & Kretzschmar 2017), which has just

169   under 100,000 words.[v] Searching "æg", "ɛg", and "eɪg" in this dictionary yielded 237 words with BAG, 181

170   with BEG, and 53 with VAGUE. Between these two lists, there were 1,149 unique prevelar words, including

171   different forms of the same lexeme like *egg*, *eggs*, *egging*, and *egged*.

172   To avoid making an excessively long survey, only a subset of these 1,149 words was selected for

173   inclusion in this study. I selected 196 words that I perceived would be common enough that most people

174   would be familiar with them—with a few exceptions. Other than the nonce words and some of the most

175   infrequent words in Jones's (2015) extensive wordlist, almost every word that has been included in previous

176   prevelar studies was included in this study. I deliberately chose to include words that are often avoided in

177   wordlists like borrowings (*allegro*, *reggae*, *magnum opus*), proper nouns (*Skagway*, *Gregory*,

178   *Copenhagen*), derived forms (*eggnog*, *pregnancy*, *megabyte*), polymorphemic words (*sagging*, *begged*,

179   *plagues*), infrequent words (*octagonal*, *interregnum*, *flagrant*), and words with orthographic <ex> (*exit*,

180   *exile*, *excerpt*, *exodus*).[vi] Some of these language-internal factors are significant predictors in prevelar

181   raising; BEG-raising is more common in frequent words and less common when the /g/ is followed by a

182   sonorant, particularly if that sonorant is a liquids (Stanley 2019). As the goal for this paper is to document

183   regional patterns in prevelar raising, language-internal predictors of prevelar raising will not be considered

184   further.

185   **3.2 The Survey**

186   To gather speaker intuition of prevelar raising across North America, I used Qualtrics, a popular online

187   platform for developing and distributing surveys, to create a categorization task in the form of an online

188   survey (Figure 1). First, participants viewed an informed consent form and, by proceeding with the study,

189   acknowledged that they agreed to its terms, including not being compensated. The survey then presented

190   each of the 196 words and asked, "How do you pronounce the following highlighted vowel sounds?" For

191   each entry, the orthographic prevelar vowel was highlighted and the word was accompanied by a brief (and

192   sometimes humorous) definition. Participants were presented with five options:
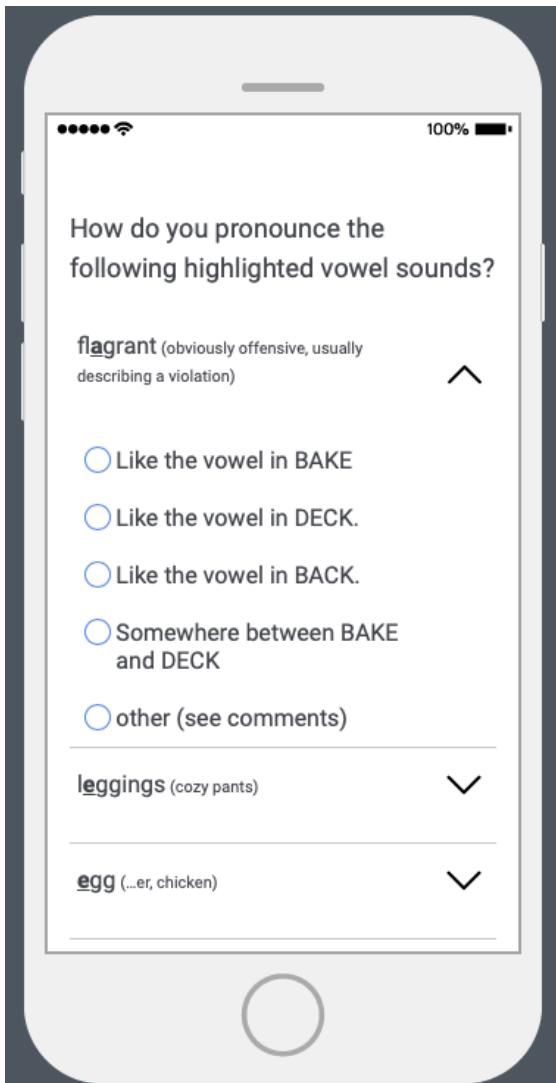
193

194   1. "Like the vowel in BAKE",[vii] which corresponded to [eɪ] and canonical VAGUE.

195   2. "Like the vowel in DECK", which corresponded to [ɛ] and canonical BEG.

196   3. "Like the vowel in BACK", which corresponded to [æ] and canonical BAG.

197   4.   "Somewhere between BAKE and DECK", which corresponded to [ẹɪ], [ɛ], or possibly [ɛ̞] to

198   account for intermediate variants.

199   5.   "other", which allows participants to indicate something else in a provided space.

200

201

202   Figure 1: The survey, as viewed on a mobile device.

203

204   Raising in the vowel space is a complex process, but such a simplification beneficial because these

205   options were presumably straightforward enough for non-linguists to understand. This simplicity comes

206   with the price of a few inherent problems though. First, self-reflection of fine phonetic detail can be

207 unreliable because people generally do not have a good intuition of their own speech (Labov 2006; Trudgill

208 1984). Furthermore, the reference words themselves are not realized the same way across the country (*e.g.*

209 speakers from the North may pronounce TRAP with a raised vowel, speakers in the South may pronounce it

210 with a diphthongized vowel, and many other speakers may pronounce it with a lowered and retracted

211 vowel).[viii] Because the purpose of this paper is to provide a general overview of prevelar raising across the

212 continent, a sufficiently large dataset should still reveal patterns through the noise. However, as is always

213 the case in survey-based phonetic studies, the results should be verified empirically with acoustic and

214 articulatory data in future research.

215   Even though less than 20% of the original 1,149 words were included in the study, a 196-question

216 survey is still quite long. To reduce fatigue, the survey was presented in blocks. The first main block

217 consisted of 69 selected words that contained enough tokens for a robust individual-level analysis. All

218 participants then proceeded to a demographic questionnaire that contained fill-in-the-blank boxes for age,

219 gender, and race/ethnicity. It also included a question that asked, "Where are you from?" and requested

220 specific cities and years of residence. After completing of this questionnaire, participants were told they

221 were done with the required portion of the survey but that they could continue to more blocks if desired.

222 Four additional optional blocks were then presented covering the remainder of the words. Within each

223 block, the same words were presented to all participants, but in a different random order each time. The

224 median time that participants took to complete the survey was 7 minutes 14 seconds if they completed just

225 the first block, and 12 minutes 34 seconds if they completed all blocks. (See the appendix for the list of

226 words in each block and a more detailed breakdown of completion times.)

227 **3.3 Survey Distribution**

228   To collect data on BEG- and BAG-raising across the continent, the sample had to have sufficient data

229 from all parts of English-speaking North America. Releasing the survey through some general distribution

230 platform or to some general random sample of North America would be inadequate for full coverage

231 because participants would presumably reflect the population distribution of the United States and Canada

232   (e.g. many Californians and few Manitobans), which was not desired for this study. Because representation

233   from many areas without regard to population density was needed, regions had to be sought out specifically.

234   Therefore, Reddit was used as the medium for distributing the survey to all areas. Reddit is a massive

235   online space where users can browse "subreddits", or communities that are about specific topics. In these

236   subreddits, users can post material related to that community, comment on the posts, and comment on those

237   comments to create a discussion. Users may also "upvote" or "downvote" posts or comments, causing it to

238   appear higher or lower in others' feeds as a result of Reddit's algorithms. Though Reddit has been used as

239   a massive corpus of online language (Baumgartner et al. 2020), this study merely uses its subreddits as a

240   tools to access specific groups of people.

241   To ensure all areas of North America were represented, I identified the most popular subreddit

242   dedicated to each US State and each Canadian province and territory (*e.g.* the subreddit for the state of

243   Georgia is called r/Georgia). I contacted the moderators for each one and requested permission to post a

244   survey. (Generally, permission is not required to post to a subreddit, but this act prevented the post from

245   being removed by the moderators and in some cases the moderators informed the community of their

246   approval, which boosted its visibility.) Only California, Colorado, Delaware, and Quebec denied the

247   request.[ix] Fortunately, because people from those areas viewed the survey in other subreddits, they are still

248   represented in the sample.

249   In some cases, moderators recommended posting the survey in city-specific subreddits, such as one

250   dedicated to Denver or cities in California; I chose not to do so because I wanted Reddit visitors from all

251   portions of the state to have an equal opportunity to view the survey. Had I selected city-specific subreddits,

252   the resulting sample would have been skewed towards urban centers at the expense of rural areas. I was not

253   able to control whether participants posted the survey to other subreddits or shared it on other social media

254   sites (and there would be no way to detect participants from these sources) so some bias inevitably may

255   have been introduced into the sample.[x]

256   When permission was granted, the survey was posted between 9am and noon local time, Monday

257   through Thursday, to take advantage of peak traffic hours. I created a dedicated username (u/dialectologist)

258 from which I explained that the survey studies "language use" in their region, without specifically

259 mentioning the linguistic variable under study.

260   In the end, 7,041 people completed at least part of the survey, yielding 578,985 total responses to survey

261 questions, mostly within two weeks of distribution (June 2018) with a small percentage coming in over the

262 following months. However, for the purposes of this study, only the 5,269 who included viable regional

263 information were analyzed (see §3.6).

264 **3.4 Data Processing**

265   Prevelar raising is not a discrete process but occurs variably along a multidimensional continuum of

266 phonetic cues like duration and vowel trajectory (Freeman 2014; Baker, Mielke & Archangeli 2008). In the

267 survey though, this complexity was simplified by providing participants with four discrete responses based

268 on similar sounding words. To convert this categorical response back into an approximate numerical value

269 of raising, participants' responses were recoded to a numeric scale from 1 to –1, as seen in Table 1.

270

271

Survey Option	Reference word	Presumed IPA	Value
1	*bake*	[e]	1
4	(intermediate)	[ẹ] or [ɛ̝]	0.5
2	*deck*	[ɛ]	0
3	*back*	[æ]	–1
5	(other)	NA	(excluded)

272   Table 1: Method for quantifying degree of raising based on participant responses.

273

274 This numeric representation of raising allow for certain quantitative analyses of this data. First, the

275 numerical values for each participant's responses were averaged, separately for BAG, BEG, and VAGUE. A

276 value close to 1 would indicate a high degree of prevelar raising across many BAG or BEG words. A score
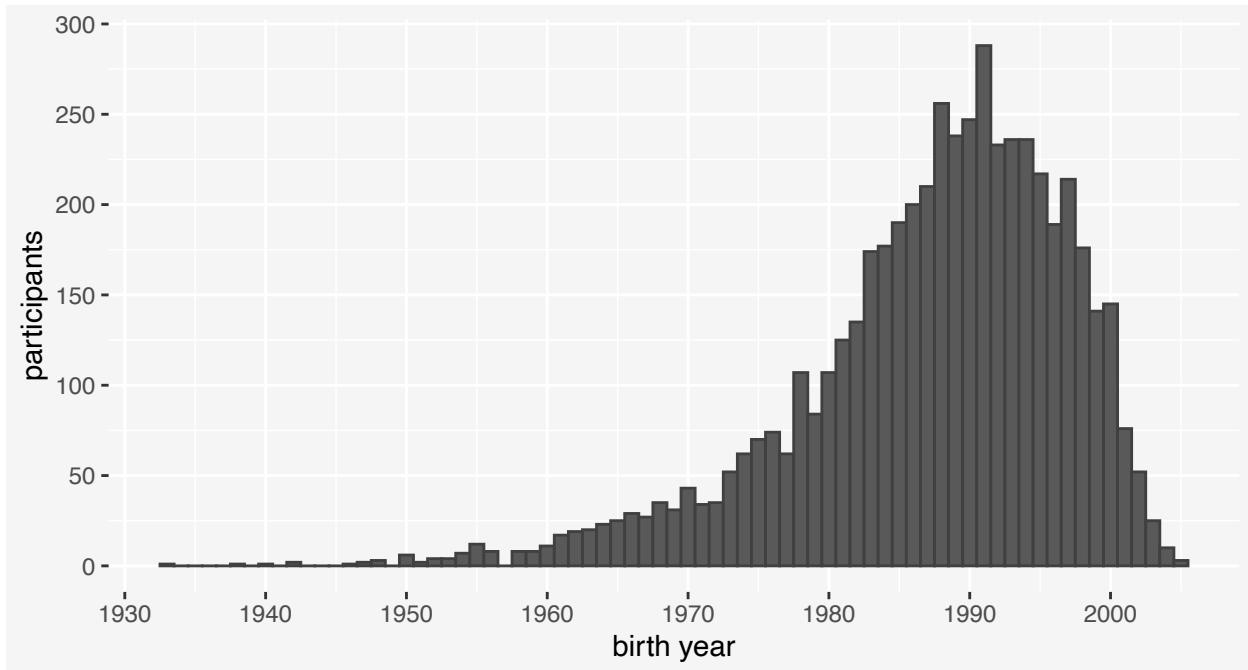
277 close to 0 for BEG or –1 for BAG would indicate very little raising. Presumably, most participants would

278 have a score close to 1 for VAGUE words since these words belong to the FACE lexical set.[xi] By using this

279 scale, I assume that participants who realize BAG as phonetically intermediate between [æ] and [ɛ] would

280 categorize some BAG words with /æ/ and others with /ɛ/, yielding a score somewhere in the middle.

281 Similarly, those who have raising in some words but not others would end up with an intermediate score as

282 well. This is an unavoidable problem in the method because two speakers with similar scores using this

283 metric may have different phonetic patterning due to the multifaceted nature of vowel shifts. This

284 quantification of the responses relies heavily on the assumption that raising is linear; acoustic measurements

285 of these same speakers' realizations of these words may reveal more nuance. Such a survey cannot fully

286 capture all patterns in prevelar raising and I encourage additional acoustic analysis to support (or refute)

287 the patterns described here.

288

289 **3.5 Participant Demographics**

290   It is important to consider the demographics of the participants in this study since they are not equally

291 balanced. Using Reddit as a medium for distribution was beneficial for ensuring that all areas of North

292 America were included. However, it introduces its own sampling bias in the other demographic factors.

293

294

295   Figure 2: Birth year distribution of the participants in this study. Created using ggplot2 (Wickham 2016).

296

297   The participants' birth year distribution was skewed in this sample, as visualized in Figure 2.[xii] The

298   majority of participants were approximately Millennial-aged, with birth years concentrated between 1980

299   and 2000. A significant minority were born between 1960 and 1980, and an even smaller number were born

300   before 1960 or after 2000. The median birth year was 1989 (29 years old when the data was collected) and

301   the average was 1987, though the spread ranged from 1933 to 2005.

302

303   Table 2: Self-reported ethnicities of the participants in this study.

Ethnicity	n	Percentage
White	4,763	90.36%
Mixed / 2 or more	196	3.72%
Hispanic	88	1.67%
Asian	47	0.89%
Native American	23	0.44%

14

African American	22	0.42%
Southeast Asian	8	0.15%
Indian	7	0.13%
Middle Eastern	6	0.11%
Pacific Islander	2	0.04%
other	24	0.46%
No response	83	1.58%

304

305   The distribution of participants' self-reported ethnicity can be found in Table 2. It is immediately

306   apparent that ethnic minorities are grossly under-represented in this sample. White participants make up

307   over 90% of the total. Other than people who identified as mixed ethnicity or who indicated two or more

308   ethnicities, no one group comprised more than 2% of the whole.

309

310   Table 3: Self-reported gender expression of the participants in this study

Gender	n	Percentage
male	3,431	65.12%
female	1,772	33.65%
nonbinary	10	0.19%
genderqueer, trans male	3 (each)	0.06%
genderfluid, other, prefer not to say	2 (each)	0.04%
agender, female genderqueer, female-ish, genderfae, male/genderfluid, male/other, male-ish, no gender, trans female, two spirit	1 (each)	0.02%
no response	33	0.63%

311

312   Finally, Table 3 shows the gender distribution of speakers in this sample. There was an overwhelming

313   majority of cisgender participants, and nearly twice as many males as females. Non-binary, transgender,

314   and many other gender identities are represented in this study, but they collectively made up less than 1%

315   of the sample.

316   When examining the age, ethnicity, and gender of the participants in this study, it is clear that this

317   sample is not representative of the general population of the United States and Canada. Instead, it skews

318   towards younger, white males and is likely reflective of Reddit users. This bias should not be surprising

319   since Reddit was used as the medium for distributing the survey. The results of this study should not be

320   interpreted as representing the speech patterns of only younger while men since minority groups are

321   represented here, but the results should be interpreted with these demographic trends in mind. Additional

322   work is required to study minority groups and regional patterns in their use of prevelar raising (Wassink

323   2016; Wassink & Hargus 2020; Clayton & Fridland 2020).

324   **3.6 Geographic Analysis and Distribution**
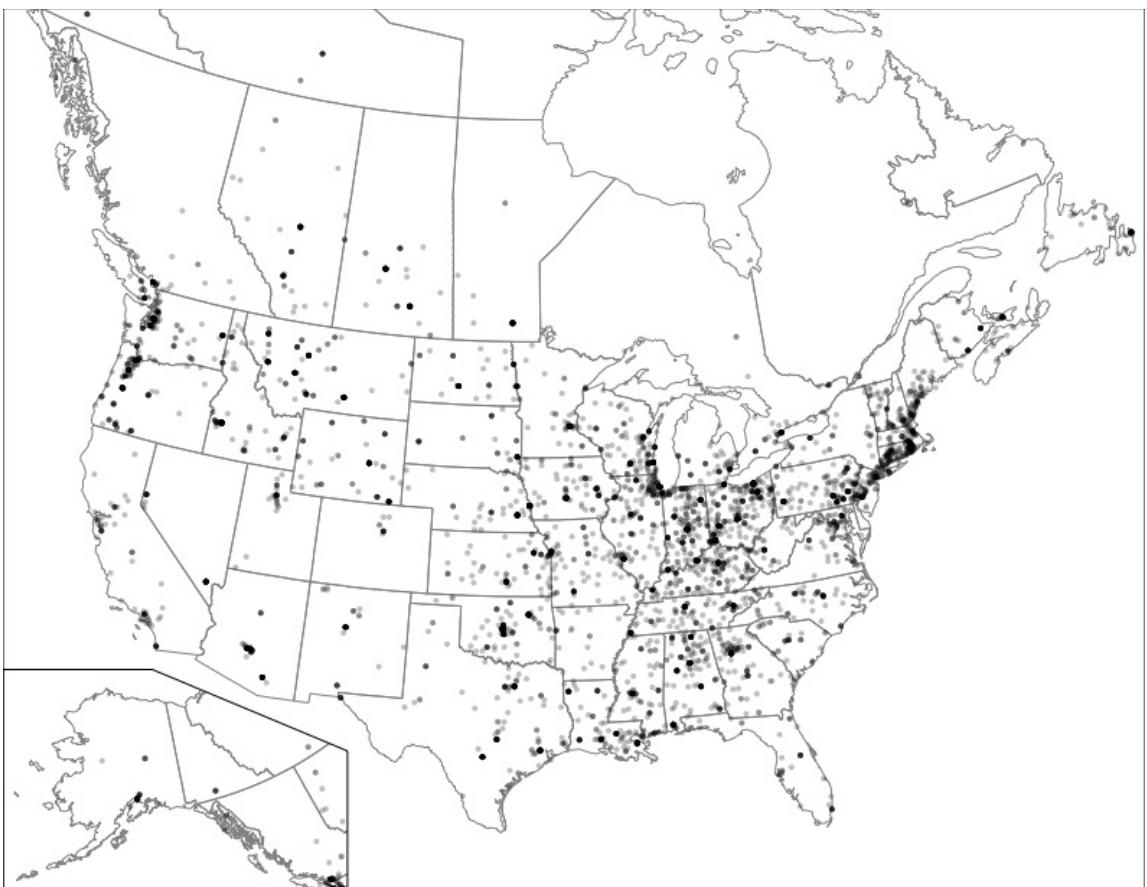
325   To process the geographic data, the first task was to assign geographic coordinates to each participant.

326   One method would have been to code them by state, based on which subreddit they saw the survey in.

327   While it is possible to generate unique links in Qualtrics and track users in that way, it would be an

328   unreliable indicator of where the participant grew up and what variety of English they might speak.

329   Furthermore, it would only provide information at the state-level, and more gradience was desired. A

330   second option was to code participants based on where their IP addresses were located. The benefit of this

331   technique is that it takes little effort and provides very detailed GPS coordinates. However, it too is an

332   unreliable indicator of where a person is from because it assumes participants were in their hometown while

333   they took the survey, which cannot be guaranteed in today's increasingly mobile society.[xiii]

334   Because these automatic methods are unreliable, participants' hometowns were coded based on their

335   responses to the fill-in-the-blank question "Where are you from?". Most participants provided detailed

336   information about which city, county, or zip code they spent their formative years. This study analyzes

337   responses from only those participants who lived in the same city from the ages of 2 to 16. That is, 5,269

338   participants, or 74.83% of the total, and 431,469 responses are analyzed in this paper. These locations were

339   geocoded and converted into coordinates using the Google Maps API via the `geocode` function in the `ggmap`

340   package (Kahle & Wickham 2013) in R (R Core Team 2018).

341   As seen in Figure 3, this sample includes at least one participant from all states,[xiv] Washington D.C.,

342   and all Canadian provinces and territories with the exception of Nunavut. Because of how the survey was

343   distributed and the desired representation from all areas, this sample is by design not a reflection of the

344   population density of North America. Places like the Northern Great Plains were over-represented: there

345   were 145 Montanans, 115 Idahoans, 79 Wyomingites, 75 Saskatchewanians, 63 North Dakotans, and 50

346   Albertans, and 29 Manitobans. Meanwhile there were just 69 Californians, who all must have seen the

347   survey through another subreddit. This increased representation in the sparsely populated areas is helpful

348   for this study because it is precisely this region that is reported to have the most prevelar raising (Boberg

349   2008). Nevertheless, this oversampling in some areas and undersampling in others necessitates that the data

350   be interpreted with care.

351
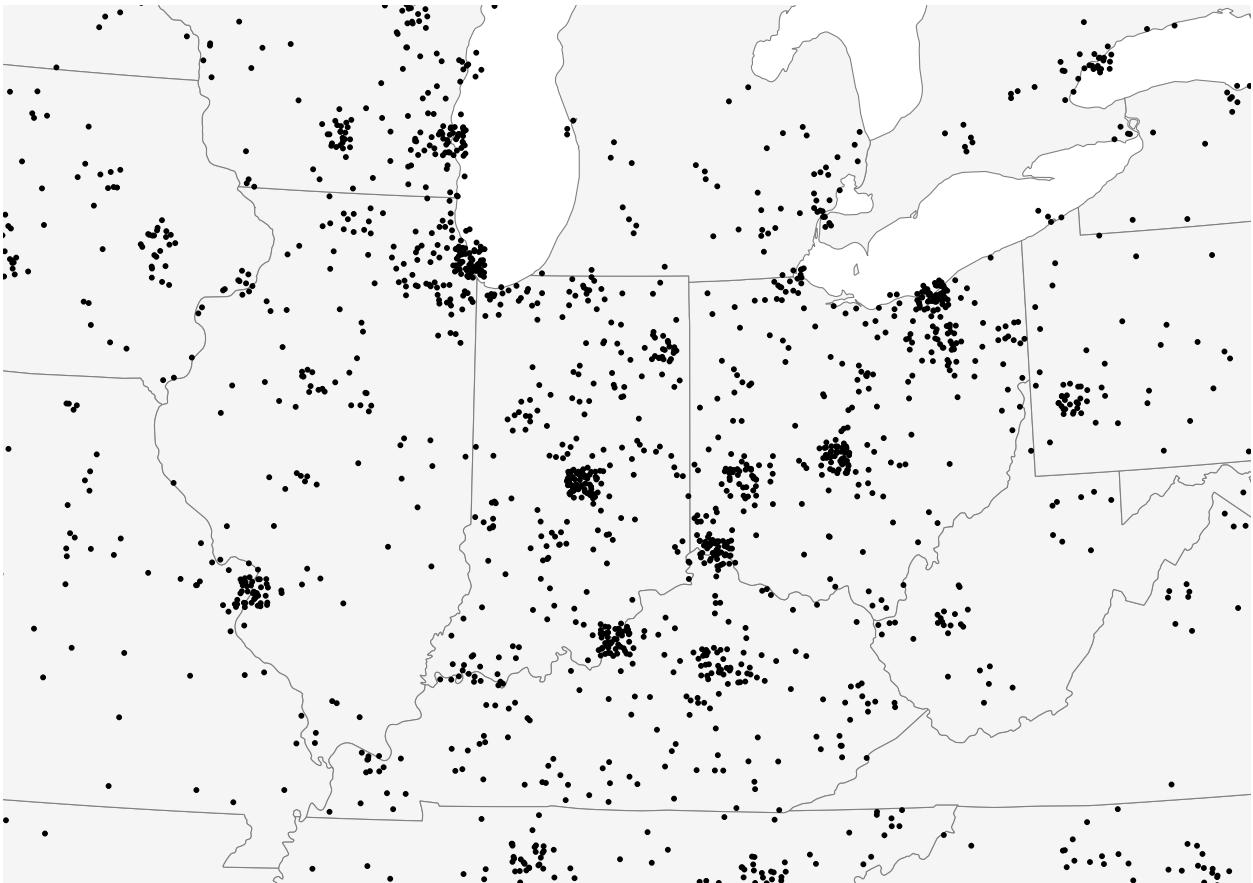
352

353 Figure 3: Distribution of the participants in this study. A small amount of jitter has been added to the

354 coordinates to allow the viewing of multiple points from the same city.

355

356

357

Figure 4: A close-up of the Midlands.

To better appreciate the amount of data represented here, Figure 4 displays a zoomed in plot focusing on the Midlands. This is the region that produced the most data: there were 376 Ohioans, 290 Hoosiers, 231 Pennsylvanians, 229 Kentuckians, 200 Illinoisans, and 186 Wisconsinites. These states' subreddits often had more comments[xv] as well. Even though this is a part of the country that is not known for its prevelar raising, these people seemed particularly interested in taking a survey that targeted this linguistic variable (or indeed, any language survey). Most of the comments revolved around other linguistic phenomena and prevelar raising was not the focus of many comments; in fact, it appears that many participants completely missed the fact that all the words in the survey had stressed vowels before /g/.

3.7 Spatial Interpolation

369

370   While data from 5,291 participants is a good amount for a dialectology study, there are inevitably gaps

371   in the data due to the vast size of North America. For example, the respondents from Nevada, Utah, and

372   Colorado were clustered in the largest urban areas, and while those states are otherwise sparsely populated,

373   there are people that live in those rural areas. Is there a way to use the data that does exist to make a guess

374   as to how areas not represented in the sample would have responded?

375   One method for spatial interpolation over such areas is ordinary kriging. This technique predicts values

376   in an unsampled pair of X-Y coordinates by considering data from neighboring points, with greater weight

377   added to closer points and less weight added to further points. *Ordinary* kriging assumes that the variable

378   being analyzed is constant within its neighborhood of points but makes no assumptions about constancy

379   between neighborhoods (as opposed to *simple kriging* which does make such assumptions).[xvi]

380   Kriging was originally used for mining and geologic engineering in the 1950s (Krige 1951), but it is

381   now widely used in a variety of disciplines. In dialectology research, kriging has been used to visualize

382   dialect areas in New England (Kim et al. 2019) and in the rest of North America (Jones 2017). It is a useful

383   tool in dialectology because it can be an objective and deterministic way to draw dialect boundaries, it

384   makes it easier to interpret a complex and noisy dataset by visualizing a smoothed underlying pattern, and

385   it shows predicted values from areas not included in a sample. For additional information on the

386   mathematics of kriging, its implementation in R, and further use in dialectology research, see Chang (2019),

387   Bivand et al. (2013), and Grieve (2013; 2018).

388   For this paper, ordinary kriging was performed using ArcGIS, a tool for the analysis and visualization

389   of spatial data. For the most part, the default parameters for the method were used: values were predicted

390   based on a neighborhood of the 12 nearest points and the semi-variogram model was spherical, meaning

391   that further points have less influence than closer points, and beyond some distance points no longer have

392   any influence. The only setting that was changed was the output cell size, which was set to 0.01 in order to

393   produce many more predicted values and therefore a higher resolution image.
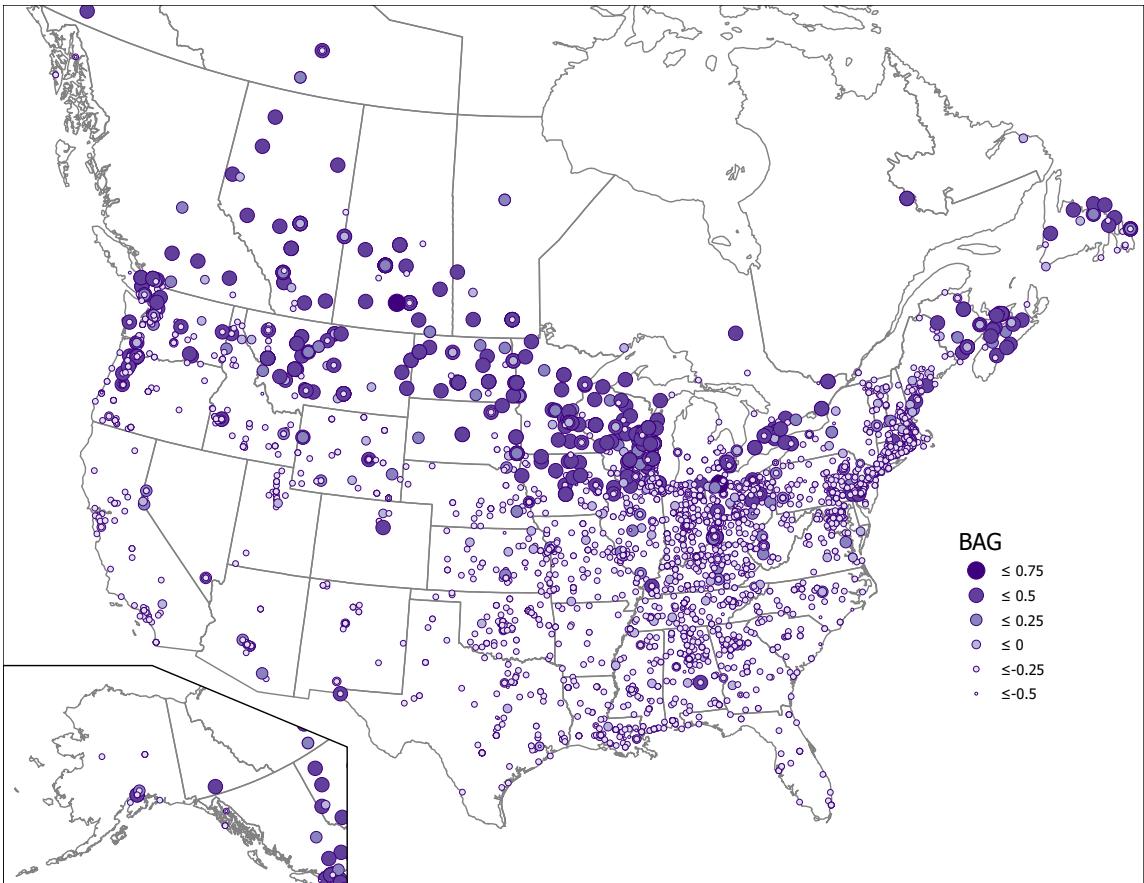
394   It should be noted that when multiple data points are located in the same point in space, the kriging

395   procedure selects one to be representative of that location. In this dataset, there were many cities that were

396   represented by multiple participants and they were rarely consistent in their BAG- and BEG-raising scores.

397   For example, the 53 participants from Chicago had a wide range of possible responses for both variables.

398   Therefore, in order to take all this variation into consideration, a small amount of jitter (on the order of a

399   few miles) was added to the coordinate data, ensuring that each participant be located in a unique point in

400   space. Using this modified dataset, the kriging procedure was then able to create an interpolated raster that

401   takes into account each participant's contribution.

402   All maps were created in ArcGIS and display geographic boundary data from IPUMS NHGIS (Manson

403   et al. 2018) using the NAD 1983 Contiguous USA Albers projection.

404   **4   Regional Patterns**

405   **4.1 BAG-raising**

406

Figure 5: BAG-raising across North America. Larger and darker dots represent more raising.
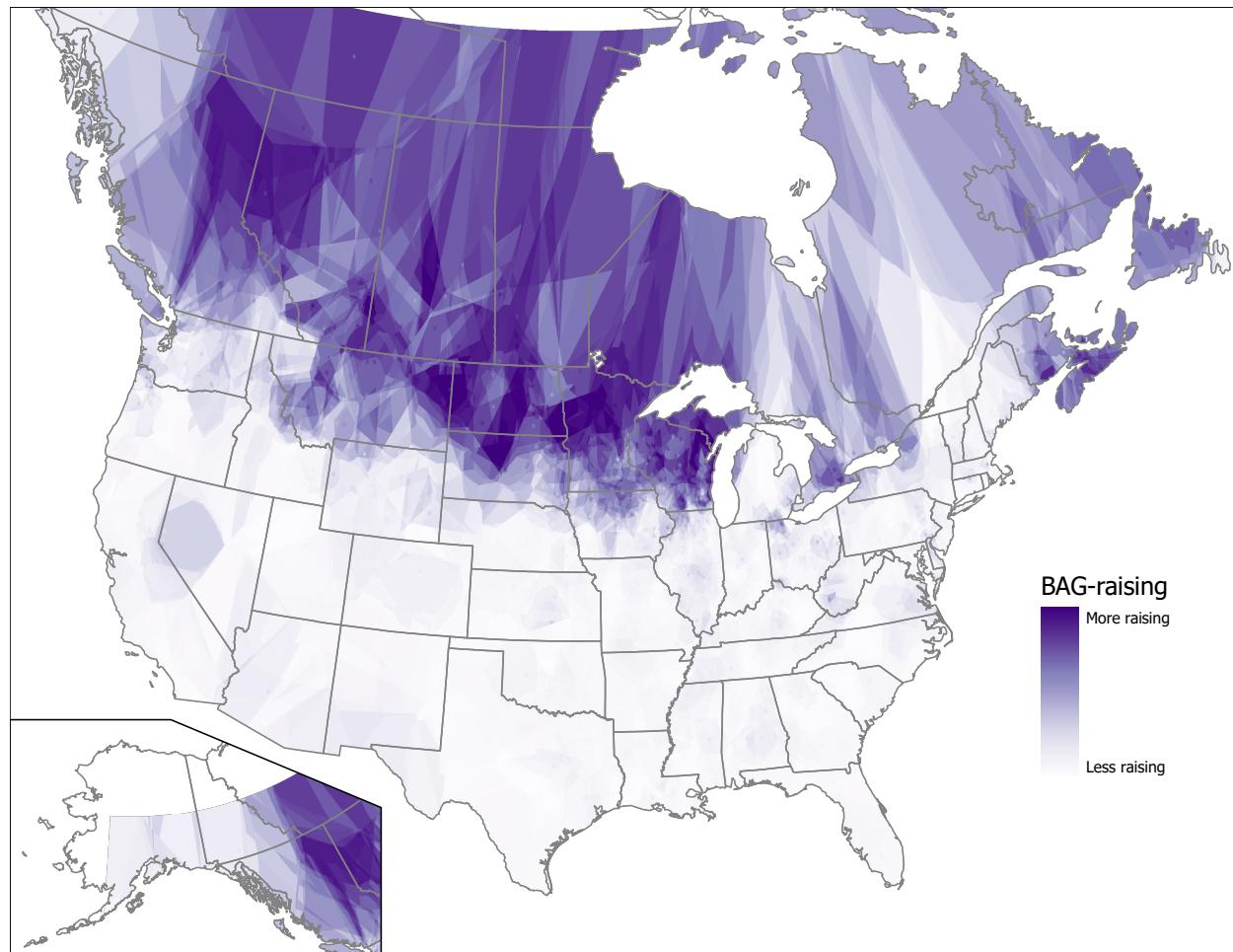
Figure 5 shows each participant plotted by hometown with the size and shading of the circle representing their degree of raising. In general, the areas where BAG-raising is reported in this sample match what has been found in previous studies. In the United States, it is most heavily concentrated in a region stretching from west of the Great Lakes to the Pacific Ocean. BAG-raising was also ubiquitous in Canada, from Vancouver to Ottawa and including the Atlantic Provinces, Yukon, and Northwest Territories.

There is variation in most regions, but the highest proportion of BAG-raisers to non-BAG-raisers is in Wisconsin, Minnesota, North Dakota, and Montana, with a modest number of BAG-raisers in the nearby states of Iowa and South Dakota as well. This heavy concentration extends into the Canadian Prairies where the majority of participants indicated a high amount of BAG-raising. Participants from Ontario and the Atlantic Provinces were also relatively homogenous in their raising. In the Pacific Northwest, while many

22

420 participants did report BAG-raising, there were many more that did not. Additionally, at least one person in

421 many other urban areas in the West had some indication of BAG-raising, suggesting some variation in these

422 areas too.

423   There were also areas with virtually no BAG-raising reported. Chief among them is the South, but in a

424 wide band stretching from Denver, across the Midlands, up through the Northeast, and all the way to Maine

425 there were relatively few BAG-raisers. In some places, the boundary between BAG-raisers and non-BAG-

426 raisers was sharp: relatively few people in the Chicago, Detroit, Buffalo areas reported BAG-raising, despite

427 their proximity to Milwaukee and Toronto.

428

430 Figure 6: Kriging analysis for BAG-raising in North American English. Darker regions represent more

431 raising. The kriging algorithm only calculated predicted values within a rectangle that includes all points,

432   so it did not include Alaska and the most northern regions of the Canadian territories since the sample did

433   not include those regions.

434
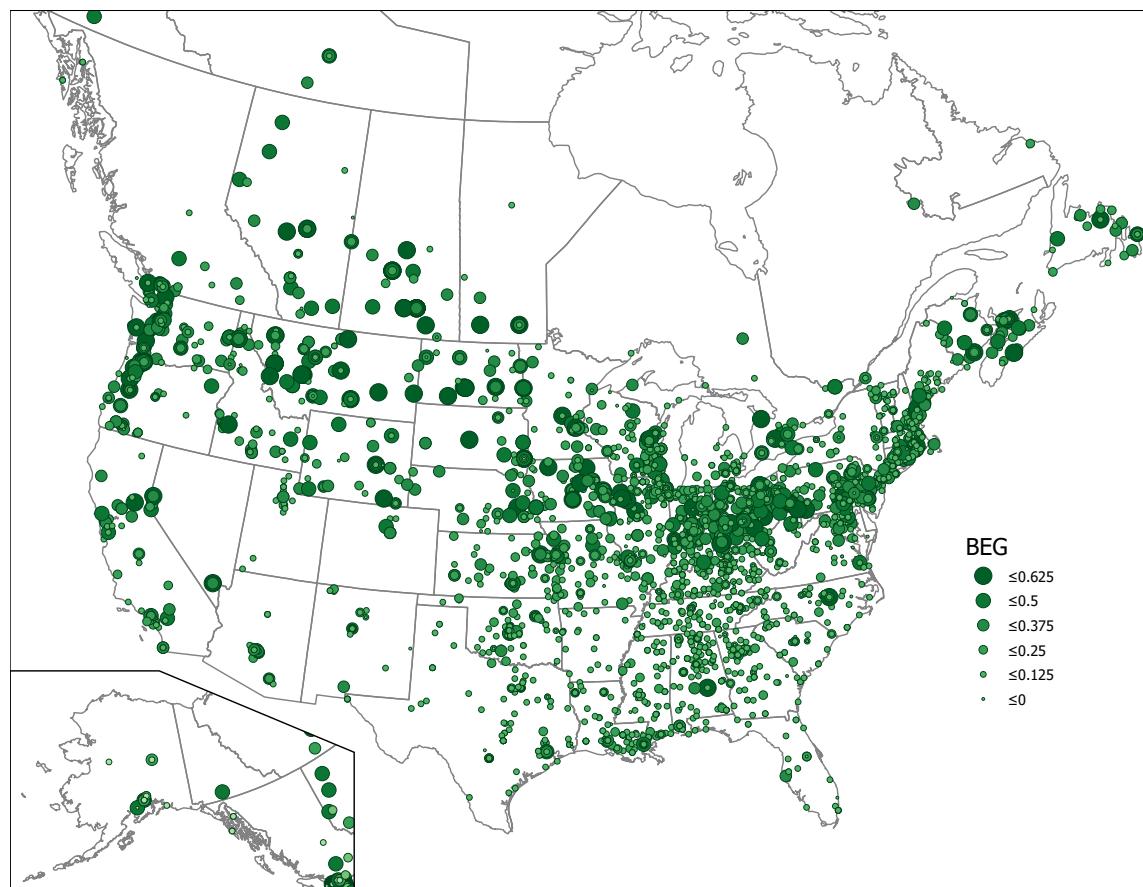
435   Figure 6 shows the output of the kriging analysis on BAG-raising, offering a smoothed version of the

436   data. This map generally agrees with Figure 5 vis-à-vis areas that have the most raising: the Upper Midwest,

437   North Dakota, the Canadian Prairies, and the Atlantic Provinces. However, the procedure has smoothed

438   over many of the individual participants that had BAG-raising when they were surrounded by many others

439   that did not. In Figure 5, when individual points were plotted, the eye is naturally drawn towards the larger

440   and more shaded points, creating a misleading impression that BAG-raising is most common than it may be

441   in some areas in the West. In Figure 6, these local outliers are washed out by the dominant pattern. This

442   map shows a clear dialect region where BAG-raising is likely to occur and not occur.

443   One potential cause for concern is that in Figure 6 the kriging predicted relatively little BAG-raising in

444   Washington State, despite it being one of the epicenters of research on prevelar raising. In acoustic analyses,

445   BAG-raising is found in men and women of all ages in Seattle and other parts of Washington (Wassink

446   2016). However, Swan (2020) and Stanley (2018) find that younger Washingtonians tend to have less

447   raising than older cohorts and that men tend to have more raising than women. Recalling back to the

448   demographics of this sample, which is overwhelmingly white, younger, and predominantly male, the

449   current study suggests that at least younger white Washingtonians *report* to have less BAG-raising, which

450   agrees with Swan's findings. Meanwhile, Swan (2020) also reports that there was more variation in Seattle

451   than Vancouver with respect to BAG-raising. This high variability is also supported in the current dataset:

452   while Figure 6 hides the many participants in Washington (and Oregon) who do in fact report a high degree

453   of raising, this variation is visible in Figure 5. In fact, a map of the variability in BAG-raising (not included

454   here) shows that there was indeed more variation in Washington's prevelar raising scores than in other

455   places, like North Dakota, which was more homogeneous. Therefore, this reported data does appear to

456   reflect the production data analyzed in previous studies.

Summarizing BAG, this data suggests that it is generally found in the same areas that have been described in previous literature as having BAG-raising: Canada, the Upper Midwest, and to some extent, the Pacific Northwest.

4.2 BEG-raising
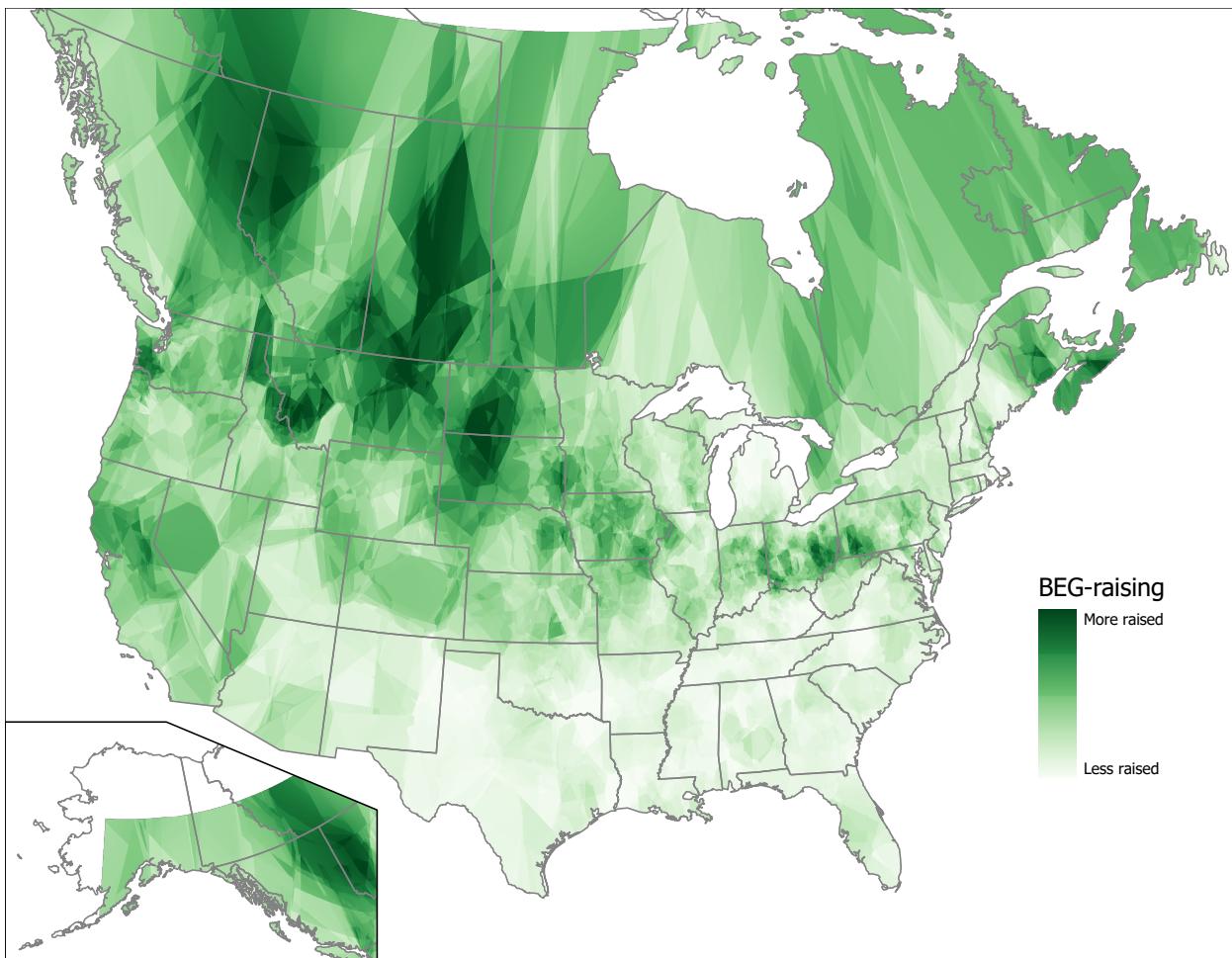
Figure 7: BEG-raising in North American English

Figure 7 displays the participants by their BEG score, illustrating that the regional distribution of BAG- and BAG- have some similarities and differences. There was a fair amount of overlap with BAG since a high amount of BEG-raising was reported the Northern Great Plains and the Atlantic Provinces. In other areas of the West, BEG is more consistently raised lower along the Pacific Coast in Oregon, and many Californians had a high degree of BEG-raising. Many more of the Idahoans, Utahns, Coloradans, and Alaskans have

470 raising. There was also higher proportion of BEG-raisers in the Midlands. In fact, the region from Denver

471 to Maine, where relatively little BAG-raising was reported, has a modest number of BEG-raisers, especially

472 in Indiana and Ohio. Similar to BAG though, there were relatively few BEG-raisers in the South.

473

Figure 8: Kriging analysis for BEG-raising in North American English

476

477   To complement the raw data, Figure 8 shows the output of the kriging function on the BEG-raising

478 scores. Again, the Canadian Prairies, the Dakotas, Montana, and the Atlantic Provinces stand out as areas

479 with a high concentration of BEG-raisers. However, what the kriging highlights that was not apparent in

480 Figure 7 are isolated hotspots of BEG-raising in areas near Pittsburgh, eastern Ohio, Cincinnati, California's

481 Central Valley (and extending into Nevada), southwest Washington, and northern Idaho. With the exception

482 of Reno, these are areas where, to my knowledge, BEG-raising has not been reported in previous literature.

483   Summarizing BEG-raising, this data suggests that it can be found to some extent everywhere but the

484 South. It is most prevalent in the Northern Great Plains, but also common in the West and Midlands, with

485 pockets of increased concentration around Ohio, California, and Washington.

486 **4.3 The independence of BAG-raising and BEG-raising**

487   When comparing Figures 8 and 10, it is apparent that there are regions where both BEG-raising and

488 BAG-raising are found, but there are also other areas where only one was robust and not the other. For

489 example, BAG-raising was especially prevalent in Minnesota and Wisconsin, but BEG-raising was not. While

490 there does appear to be a fair number of people who reported BEG-raising in that region, raised variants

491 were not in the majority. This suggests that the Upper Midwest is a dialect area that has BAG-raising without

492 BEG-raising. Furthermore, the "hotspots" of BEG-raising all occurred where BAG-raising was not found.

493   To illustrate these dialect areas where one vowel raises without the other, a composite score that

494 combines both prevelar raising types was calculated out of the two kriging outputs. The new score is simply

495 the result of subtracting the predicted BEG-raising score from the BAG-raising score, after rescaling the BAG-

496 raising score to a scale of 0 to 1 to match the BEG-raising range. In other words, it is the result of overlaying

497 and averaging Figure 6 and Figure 8 (cf. Kim et al. 2019: 182 Figure 30). This method produces a new

498 score that quantifies the *type* of prevelar raising prevalent in the area. Higher scores (near 1) are the result

499 of a high amount of BAG-raising but very little BEG-raising. Lower scores (near -1) suggest little BAG-raising

500 but a high degree of BEG-raising. Scores near zero are ambiguous: they indicate equal amounts of BAG-

501 raising and BEG-raising. Whether a zero means no raising whatsoever or extreme amounts of both can only

502 be interpreted with the help of other maps.
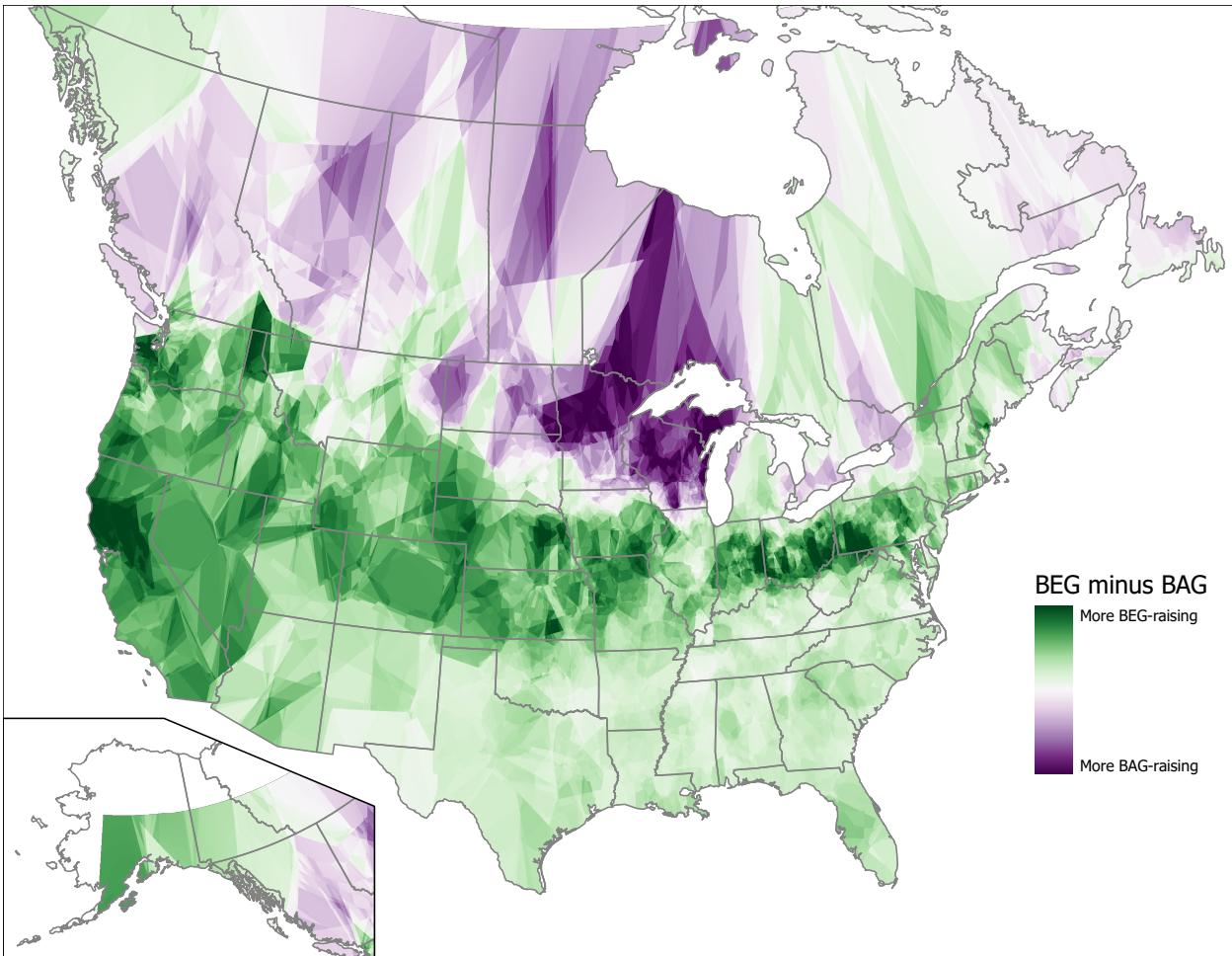
Figure 9: Difference between the interpolated BAG-raising and BEG-raising scores.

This combined prevelar raising score across North America is visualized in Figure 9. Just as predicted, Minnesota, Wisconsin, and eastern North Dakota (in the vicinity of Fargo) were areas that primarily had BAG-raising and relatively little BEG-raising. There is also some indication of this type of prevelar raising in the Maritime Provinces. The map also indicates that Michigan's Upper Peninsula and Eastern Ontario are areas where only BAG-raising is found. However, there were very few participants from these regions (see Figure 3), so this northward extension is simply an effect of the kriging algorithm interpolating the unknown area with the nearest known data. Whether people in these areas actually have prevelar raising is unknown given this sample.

514   Conversely, there is a large region where BEG-raising was reported but not BAG-raising. Starting in

515   Pennsylvania, it extends westward, gradually widening until the Rocky Mountains at which point it spreads

516   to the majority of the American Pacific Coast. All the previously mentioned "hotspots" of BEG-raising, are

517   located in this band. It is notable that this region coincides closely to settlement patterns as well as the

518   Midlands and Western dialect boundaries in the *Atlas of North American English* (Labov, Ash & Boberg

519   2006). One part of the Midlands with less raising is Illinois, which corresponds to the St. Louis Corridor,

520   supporting its linguistic similarity to the North. In Canada, some areas are more BEG-raising–dominant

521   (Saskatchewan, Quebec, parts of Ontario) while others are more BAG-raising–dominant (British Columbia,

522   Alberta, Manitoba). Meanwhile, Alaska participants reported more BEG-raising, supporting its inclusion

523   into the Western dialect area (Bowie 2020). Though BEG-raising without BAG-raising has been described

524   as "a pattern that seems to be unique to Nevada" (Clayton & Fridland 2020: 52), this sample suggests that

525   it is far more widespread than any one state or region.

526   Figure 9 demonstrates that while prevelar raising sometimes follows established dialect boundaries, it

527   crosses others. In particular, the boundary between Canada and the Inland North is blurred as speakers in

528   both areas raise both BAG and BEG. However, between BAG-raised Upper Midwest and BEG-raised

529   Midlands, it is still sharp. The characteristic heterogeneity of the West is retained, as some areas have many

530   participants with prevelar raising and others have none. Even though many people with prevelar raising

531   were found in Wisconsin, there were virtually none in Michigan and New York, illustrating a divide in the

532   Inland North. There is some variation in New England and the Atlantic Provinces with some hint of regional

533   patterning (like BEG-raising dominant in Irish-settled Newfoundland and BAG-raising dominant elsewhere)

534   but these were minority speech patterns. Finally, as was evident in previous maps, the South and even the

535   southwestern states were resistant to prevelar raising, making this phenomenon exclusive to Canada and

536   the northern half of the United States (including the portion of Alaska included in the sample).

5 Discussion

This paper presents two hypotheses regarding regional patterns in prevelar raising. First, prevelar raising is more widespread than has been reported so far. The data presented in this paper is evidence to support this claim. In prior studies, BAG-raising has been described in areas including the Upper Midwest, Canada, and the Pacific Northwest. This survey-based data supports these same regions, with the large number of participants adding some clarity to the boundaries of the existing dialect area.

However, there has been less work conducted on BEG-raising, most of which has been in areas known to exhibit BAG-raising. The data presented in this paper, which includes participants from most of North America, shows that BEG-raising is found in areas outside of the Pacific Northwest. It is scattered across most of the Western States, the Midlands, Canada, and the Martine Provinces. In fact, the only region where BEG-raising was not prevalent was the South.

The other hypothesis that this data supports is that BEG- and BAG raising do not always occur together. Their coexistence was never really implied in the limited research on BEG-raising, and evidence for one without the other has already been found in Vancouver (Mellesmoen 2018) and Reno (Gunter, Clayton & Fridland 2017). However, this data suggests that entire regions may be characterized as having just one vowel raised, like the Upper Midwest and BAG-raising or Ohio and BEG-raising. Meanwhile people in Northern Great Plains reported both vowels as raised.

5.1 Vowel Shifts

This paper presents findings from data based on a survey of vowel class membership. As was mentioned previously, one potential caveat is that the reference words (*bake, deck*, and *back*) are not pronounced the same across North America. In fact, those three vowels (FACE, DRESS, and TRAP) are involved in various shifts occurring in North American English. Do the patterns reported in this study simply reflect regional vowel shifts rather than prevelar raising? Because acoustic data was not gathered from these people, there is no way of establishing whether other vowel changes are present in their speech and the possibility that

562   these results are the result of other vowel shifts cannot be ruled out. And as one anonymous reviewer points

563   out, the survey technically does not gather information about raising but rather word class membership and

564   what may be interpreted as a raised prevelar vowel may in fact be the result of a lowered reference vowel.

565   However, by analyzing each of the major vowel shifts in turn and comparing geographic patterns of acoustic

566   descriptions with the reported data here, I argue that the vowel shifts likely had little influence on the overall

567   findings.

568   The general lack of reported prevelar raising in the South could be interpreted as evidence of some

569   confounding factor. Speakers with the Southern Vowel Shift (Labov, Yaeger & Steiner 1972; Labov, Ash

570   & Boberg 2006; Thomas 2003) might realize the nuclei of *bake* lower and *deck* higher than speakers without

571   the shift. If these southerners did have some amount of prevelar raising, perhaps their hypothetically raised

572   BEG is indeed perceptually closer to the raised DRESS in *deck* ([eə]) than to the lowered FACE in *bake* ([ɛɪ]).

573   However, it has been found that that the Southern Vowel Shift is absent in some younger southerners

574   (Dodsworth & Kohn 2012; Stanley In press), which is the same demographic that this study represents, so

575   it may be the case that many of the participants in this sample do not have the Southern Vowel Shift. If

576   reported prevelar raising were entirely determined by presence of shifted vowels in accordance to the

577   Southern Vowel Shift, we would expect to see some amount of variation in reported prevelar raising

578   patterning the variation in the Southern Vowel Shift. Furthermore, we would also expect to find less

579   prevelar raising in areas of the South where DRESS and FACE are less affected by the shift (*i.e.* Georgia,

580   Florida, and Southeastern Texas; Labov, Ash, & Boberg 2006:249). However, no such patterning exists in

581   the reported prevelar data presented here.

582   Another relevant shift is the Northern Cities Shift (Labov, Ash & Boberg 2006) which raises TRAP and

583   either lowers or centralizes DRESS. These shifted phonemes may cause the prevelar words to appear to be

584   more or less raised when in fact it is simply a manifestation of other regional patterns. A speaker with the

585   Northern Cities Shift may pronounce *back* with a phonetically higher vowel than in *deck*, so if their raised

586   BAG or BEG indeed sounds more like *back* than *deck*, they would end up with a score closer to –1. If reported

587   prevelar raising were strongly influenced by the Northern Cities Shift, we would find that regions where

588   this shift occurs would have low scores. There were indeed 95 participants who had a negative score for

589   BEG, indicating more association with *back* than *deck*, which may be evidence for this influence. However,

590   these participants came from 42 different states/provinces and were not regionally clustered in any way.

591   Furthermore, regions with low BAG-raising scores were Michigan and Upstate New York, two regions

592   where the Northern Cities Shift is receding (Nesbitt, Wagner & Mason 2019; Thiel & Dinkin 2017).

593   Therefore, it does not seem likely that the Northern Cities Shift is a strong correlate of reported prevelar

594   raising in this data.

595   One final shift that is prominent across North America is the Low-Back-Merger Shift (Hinton et al.

596   1987; Clarke, Elms & Youssef 1995; Becker 2019), which lowers DRESS and lowers and centralizes TRAP.

597   In speakers with advanced versions of this shift, BAG in its historical position would be closer to *deck* than

598   *back*. In fact, in San Francisco, Cardoso et al. (2016) find that while BAG is not necessarily raised, it appears

599   to lag behind the lowering and centralizing that is affecting TRAP in other phonological contexts. If reported

600   prevelar raising were correlated with the Low-Back-Merger Shift, we may find that areas where this shift

601   is widespread have more reported BAG- and BEG-raising. While the current study shows that both BAG and

602   BEG are reportedly raised in Canada, BAG at least is not reported to be raised in California where the shift

603   is robust. BEG-raising is reported in California, and it may be that DRESS has lowered so much that BEG in

604   its historical position is indeed closer to the FACE vowel in *bake* than the shifted vowel in *deck*. However,

605   in speakers with the Low-Back-Merger Shift, the amount of shifting found in TRAP typically outstrips that

606   of DRESS (Becker 2019), so if there was any correlation between reported prevelar raising, it would be more

607   visible in BAG-raising rather than BEG-raising.

608   Further evidence that the results here are not simply a reflection of regional vowel shifts can be found

609   in their geographic patterning. The regions where BAG-raising was reported the most often in this study

610   very closely align with areas where BAG-raising has been documented using acoustic data, including the

611   Upper Midwest and Canada. Furthermore, within these regions where BAG-raising is known, the elsewhere

612   allophones of these vowels are undergoing distinct changes: TRAP is raised in the Upper Midwest as a result

613   of the Northern Cities Shift (Labov, Ash & Boberg 2006) while it is lowering and retracting in Canada

614   because of the Low-Back-Merger Shift (Boberg 2019). Nevertheless, BAG-raising is reported in both

615   regions.

616   Using reference words is not a perfect system, and additional acoustic data is needed to verify the

617   regional patterns presented here, but the large amount of data and the clear regional trends show patterns

618   that are most likely manifest in spoken language, regardless of the presence of a regional vowel shift.

619

620   **5.2 Is the longer wordlist justified?**
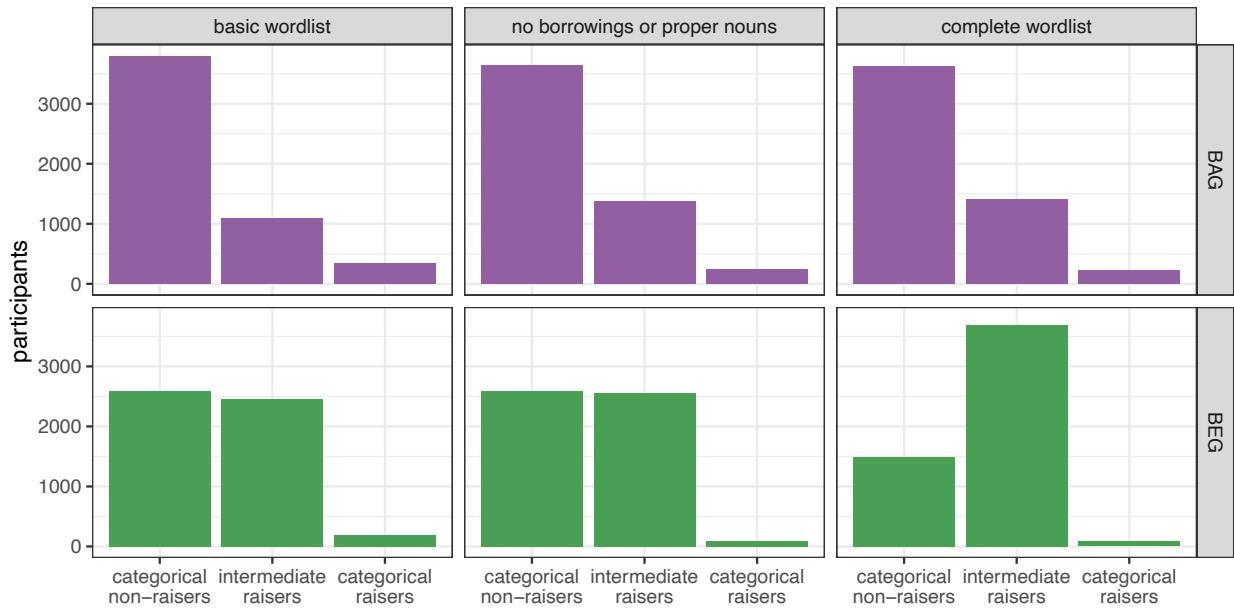
621   As mentioned in §2, nearly all past research on prevelar raising is based on a relatively small set of

622   words. For BAG, studies typically choose a subset of *bag*, *Bagdad*, *brag*, *crag*, *drag(ging)*, *dragon*, *flag*,

623   *gag*, *hag*, *haggle*, *lag*, *magnet*, *nag*, *pragmatic*, *sag*, *stag*, *tag*, and *(zig)zag*. Meanwhile, for BEG, the usual

624   words to study are *beg*, *egg*, *keg*, *leg*, *Meg*, *peg*, *leggings*, *regular*, *negative*, and *Peggy's*. This study has

625   expanded the wordlist to include a much larger set of words—the question is whether this expanded word

626   list was justified.

627   Part of the inclusion of a longer list is based on my own speech. I have prevelar raising in most BEG

628   words, but not all. For example, *integrity, interregnum*, and *segment* are not raised in my idiolect. Word

629   frequency may partially explain the exceptions, but I do have raising in *renege*, which is quite infrequent.

630   Surrounding phonological context may explain some of it too, though I raise the vowel in *negligible* but

631   not *negligent*. Perhaps most interestingly, I raise the vowel in *peg* but not *JPEG*. Had I been included in a

632   previous study and read only the small set of words typically included in a BEG-raising study, the researcher

633   would conclude that I categorically raise BEG when, in reality, I do not. It is the words that are typically *not*

634   included in these studies that are exceptions to my idiolect's raising.

635   To see whether the expanded wordlist was justified or if it introduced redundancy and extra

636   complications into the analysis, the full dataset was compared to two subsets of the data. One subset

637   includes just the relatively few words that are normally included in previous studies. The other subset strikes

638   a middle ground and includes all words except borrowings or proper nouns since they may not pattern like

other words in their lexical class (cf. Hall-Lew, Friskney & Scobbie 2017). The average BEG and BAG scores were then recalculated for each person for each subset of words, yielding six numbers: one for each of the sets of words (basic, non-proper, all) for each vowel class (BAG and BEG). The distributions of these numbers are visualized in Figure 10.


Figure 10: Amount of raising by vowel and subset. Categorical non-raisers are those who scored less than –0.90 for BAG and 0.05 for BEG while categorical raisers are those that scored greater than 0.9 for BAG and 0.95 for BEG; in other words, anybody scoring in the top or bottom 5% of the possible range of values is considered "categorical." Intermediate raisers are those with anything in between.

Beginning with BAG (the top three panels), there was admittedly little difference between the subsets. The number of intermediate raisers is slightly lower in the basic wordlist compared to the other sets of words, but it is not a large change. This suggests that reported BAG-raising is consistent across the lexicon, whether it be in common words (like *bag* or *flag*), proper nouns or borrowings (like *Skagway* or *magnum opus*) or infrequent words (like *pentagonal* or *magnolia*). Therefore, previous researchers may be justified

655   in choosing the words that they have been using since there are diminishing returns when more words are

656   elicited.

657   On the other hand, the heights of the bars on the three bottom panels, which represent BEG-raising

658   across the three subsets of the data, are markedly different. Specifically, there is a drastic increase in the

659   number of people classified as "intermediate raisers" when the full wordlist is employed. In other words,

660   people that do not exhibit BEG raising in the words that are typically studied (like *beg* and *legs*) or even

661   proper nouns and borrowings (like *Greg* and *oregano*) may in fact exhibit raising in infrequent words (like

662   *negligence* or *interregnum*)—which is opposite the pattern in my own speech.[xvii] This suggests that BEG-

663   raising has not fully diffused across the lexicon and supports the lexically-driven tendencies found in

664   acoustic data (Gunter, Clayton & Fridland 2017). Analyzing only the most common words would

665   overestimate the amount of BEG-raising that may actually exist in a person's speech.

666   Besides the anecdotal distribution in my speech and the patterns this dataset reveals, it is important that

667   studies on the same phenomena not limit themselves to the same restricted set of words. Replicating and

668   making direct comparison with existing literature is useful, but it is also important to uncover potential

669   nuances that previous studies have inadvertently overlooked. To draw a parallel in traditional dialectology,

670   it is good to study nonmobile, older, rural men ("NORMs") since their speech is expected to reveal

671   traditional, conservative speech patterns; however, sociolinguistic studies in the past 70 years have shown

672   that is also important to study mobile, younger, and urban people of all genders, ethnicities, and social

673   classes to fully describe the variation that exists. Similarly, a study cannot sample words with the least

674   amount of phonological "baggage" (like potential effects of syllable count, syllable structure, word

675   frequency, word origin, and surrounding consonants) and then extrapolate the findings to all words in that

676   lexical set. Some of the variation may be in the baggage itself. So while the results in this study may not be

677   perfectly comparable to previous studies on prevelar raising due to the different selection of words, I argue

678   that this study provides a glimpse at the fuller view that other work has overlooked precisely because of the

679   larger set of words.

6   Conclusion

One purpose of this paper was to show that prevelar raising is more widespread than has been previously reported. Using survey data collected from more than 5,000 participants, this study has provided a more complete picture of prevelar raising across the United States and Canada. The area where BAG-raising was most often reported coincided with where it has been previously documented using acoustic data. However, BEG-raising was more widespread, occurring in most areas except for the South. When overlaying the two distributions, this data suggests that while both BAG- and BEG-raising are found in Montana and the Canadian Prairies, BAG-raising by itself is dominant in the Upper Midwest and BEG-raising by itself is found in the Midlands and scattered across the West. In other words, BAG-raising and BEG-raising may be manifestations of the same motivating factor, but their geographic distributions indicate that they are not co-dependent and that one can occur without the other.

Another purpose of this study was to identify how widespread BAG- and BEG-raising are across the lexicon. Using an extensive wordlist with dozens of tokens of each vowel, this study finds that while BAG-raising may be uniform across the entire lexical class, BEG-raising is reported to occur more often in the shorter, frequent words than in the infrequent ones. These results suggest that while a reasonably accurate picture of BAG-raising can be captured using a smaller wordlist, the full extent of BEG-raising can only be uncovered using a longer wordlist because for many people raising has not yet spread to the infrequent lexical items.

This study also shows that websites like Reddit can be useful tools for dialectology in the 21st Century. Taking advantage of pre-existing social spaces that coincide with the populations of interest proved very effective in this study. Such a technique allowed rural regions, which do not normally play prominent roles in recent dialect studies, to be included as a significant part of this sample. In fact, it was these regions that were the most interesting in regard to prevelar raising. The same method can be used to target minority and other marginalized populations by identifying online spaces where these people congregate.

Because of the limitations of the survey and the medium of distribution, these findings should be backed up with acoustic data from a more demographically balanced group. The constraints of the survey

36

706 necessarily limit the reduction of a complex continuous variable to a categorical one entirely based on
707 introspection. Furthermore, because Reddit was used as a recruitment tool, this sample likely reflects the
708 userbase for that website (younger White males), which is not reflective of the population of interest
709 (English speakers in North America). Additional work on regional patterns in prevelar raising should
710 address these gaps and identify patterns in minority groups.

711 The purpose of this study was to identify potentially new areas where prevelar raising may be
712 occurring. The data presented here suggest that researchers in all parts of North America (except for the
713 South) may find some amount of prevelar raising in their speakers. In fact, because there was variation in
714 all regions, acoustic analysis may help uncover sociolinguistic patterns that correlate with BAG- and BEG-
715 raising. Labov stated that "[i]n listening to everyday speech, we tend to hear only those linguistic features
716 that have already been described, and it takes a major effort to hear the new variables that are being
717 generated in the speech community" (2006: 27). Now that prevelar raising has been described to some
718 extent in all of North America, additional researchers may hear it more.

719 **7   Acknowledgements**

724

725

726 **8   Appendix**

727 The following 69 words appeared in the initial block of the survey, which all participants viewed. 4,065
728 participants (77%) completed this block (and the demographics block) but did not progress further. Their
729 median completion time was 7 minutes 17 seconds.

730

agony	flagrant	maggot	peg	segment
agriculture	fragment	magma	Pegasus	segregate
bag	fragrance	magnet	Peggy	segue
bagel	Greg	magnify	plague	shaggy
Baghdad	integrity	Meg	pregnant	stegosaur(us)
beg	interregnum	mega	Prego	straggler
Craig	jaguar	Megan	protagonist	tag
diagonal	JPEG	Montenegro	rag	vagabond
dragon	keg	negative	reggae	vagrant
drags	Las Vegas	negligent	regular	vague
dregs	leg	nutmeg	regulate	wagon
Egbert	legacy	omega	Regulus	Winnebago
egg	leggings	oregano	sag	Winnipeg
flag	Lego	pagan	San Diego	

731

732   The following 29 words then appeared in the next block. There were 615 participants (12%) who completed

733   this block but did not progress any further. Their median completion time was 9 minutes 51 seconds.

aggravate	Copenhagen	Hague	nag	scallywag
antagonist	dagger	jagged	Niagara	shag
baggage	flagstaff	magazine	Pythagorean	Skagway
bodega	gag	Magna Carta	ragged	stag
brag	Gregory	magnum opus	Reagan	swagger
coagulate	hag	Mary Magdalene	rutabaga	

734

735   Then, these 34 words appeared in another block. There were 265 participants (5%) who completed this

736   block but did not progress any further. Their median completion time was 13 minutes 21 seconds.

aggregate	dragnet	lag	octagonal	swag
allegro	Fagan	lolligag	Paganini	Trinidad and Tobago
baggy	fragrant	Maggie	pentagonal	Viagra
crag	gaggle	magnificent	pragmatic	Volkswagen
Diego	Gregg	magnitude	stagger	Wagner
dishrag	haggard	magnolia	stagnant	zigzag
dogtag	hexagonal	Magnus	stagnate	

737

738 This block, which includes 53 polymorphemic words, appeared next. There were 59 participants (1%) who

739 completed this block but did not progress any further. Their median completion time was 14 minutes 58

740 seconds.

bagged	*brags*	*irregular*	*megaphone*	*plaguing*	*tagged*
bagging	*eggnog*	*kegs*	*nagging*	*pregnancy*	*tagging*
bags	*eggplant*	*lagged*	*nags*	*rags*	*tags*
begged	*flagged*	*lagging*	*negligible*	*regularize*	*wagged*
begging	*flagging*	*lags*	*pegboard*	*sagging*	*wagging*
begs	*flags*	*legging*	*pegging*	*sags*	*zigzagged*
bootleg	*gagged*	*leghorn*	*pegs*	*snagging*	*zigzagging*
bragged	*gagging*	*legs*	*plagued*	*snags*	*zigzags*
bragging	*gags*	*megabyte*	*plagues*	*stags*	

741

742 Finally, the last block included these 11 words with orthographic <ex>. Note that participants were also

743 asked whether they voiced the following consonant cluster and responses where voicing was indicated were

744 retained. There were 260 participants (5%) who viewed this block and therefore completed the entire

745 survey. Their median completion time was 17 minutes 53 seconds.

eczema	*exanthema*	*exhortation*	*exigency*	*existentialism*	*exodus*
exaltation	*excerpt*	*exigence*	*exile*	*exit*	

746

747

---

[i] Though the environment that triggers this raising in voiced velars—that is, [g] and [ŋ]—this paper focuses exclusively on TRAP and DRESS before [g].

[ii] Anecdotally, in my own St. Louis–based idiolect, I do not have BAG-raising but I do raise most BEG words. To my knowledge, BEG-raising has not been studied in the American Midlands so prior to this study I did not know if my speech patterns were typical of that region. My own speech patterns combined with this regional gap in the literature were the motivation to do this study.

[iii] Researchers who study certain morphosyntactic variables have had the same issues and have also resorted to alternative approaches to data collection.

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

v I am grateful to Bill Kretzschmar for providing me with a searchable version of this dictionary.

vi In the case of words with orthographic <x>, participants were also asked whether they pronounced the <x> "like 'gz' or like 'ks'"; only responses that indicated voiced consonant clusters were retained for analysis. These words were included in the last block of the survey and relatively few people saw them so these words only make up 0.28% of the total sample.

vii The reference words *bake*, *deck*, and *back* were used because they are common words that end in a voiceless velar stop and act as a way to contrast the voiced stop in the target words.

viii A potential solution to this issue would have been to include audio as a part of the survey and have participants rate their pronunciations relative to those recordings. I chose not to use this method because I wanted the survey to be as brief and accessible as possible, and the additional task of listening would have slowed people down in an already tedious survey.

ix The California subreddit has thousands of subscribers and the moderators kindly explained that if surveys were allowed, the subreddit would quickly turn into nothing but surveys to take advantage of the large amount of viewers. The moderators for the Quebec subreddit denied the request because they wanted to retain the primarily French-based nature of the community.

x I do know that the survey was cross-posted to r/Cincinnati and other Ohioans shared it elsewhere on social media, which may explain the large number of participants from Ohio.

xi Though Freeman (2014) reports VAGUE-lowering among speakers with prevelar raising, there was relatively little evidence of that in this dataset.

xii Recall that age was a fill-in-the-blank box, so this figure does not show the five participants that put a vague answer like, "late twenties".

xiii In fact, approximately half the participants took the survey in a different state from the one they reported growing up in.

xiv Hawaii was excluded from this study because it is not contiguous with the rest of North America.

xv There were hundreds of comments on these posts that included metalinguistic commentary; I hope to incorporate them into a future analysis of prevelar raising.

^{xvi} An alternative technique is EBK Regression Prediction. This takes ordinary kriging to the next level by incorporating independent variables into the model. Since this independent variable has to be in the form of a raster, I could only include age, since it was the only numeric demographic variable collected. The output of this model would show prevelar raising after taking into account any age-related differences. However, since age was more or less randomly distributed across the region, the results of this method were largely the same the ordinary kriging output.

^{xvii} Incidentally, the number of categorial BEG-raisers decreases when with longer wordlists. My own speech patterns would put me as one of those people.

748

9   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: The Mountain West* (Publication of the American Dialect Society. 102), 107–138. Durham, NC: Duke University Press. DOI: 10.1215/00031283-4295299.

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

Baumgartner, Jason, Savvas Zannettou, Brian Keegan, Megan Squire & Jeremy Blackburn. 2020. The Pushshift Reddit Dataset. *Proceedings of the Fourteenth International AAAI Conference on Web and Social Media* 14. 830–839.

Becker, Kara. 2019. Introduction. In Kara Becker (ed.), *The Low-Back-Merger Shift: Uniting the Canadian Vowel Shift, the California Vowel Shift, and short front vowel shifts across North America* (Publication of the American Dialect Society 104). Durham, NC: Duke University Press.

Becker, Kara, Anna Aden, Katelyn Best & Haley Jacobson. 2016. Variation in West Coast English: The case of Oregon. In Valerie Fridland, Tyler Kendall, Betsy E. Evans & Alicia Beckford Wassink (eds.), *Speech in the Western States, Vol. 1: The Pacific Coast* (Publication of the American Dialect Society 101), 107–134. Durham, NC: Duke University Press. doi: 10.1215/00031283-3772923.

Benson, Erica J., Michael J. Fox & Jared Balkman. 2011. The bag that Scott bought: The low vowels in northwest Wisconsin. *American Speech* 86(3). 271–311. https://doi.org/10.1215/00031283-1503910.

Bivand, Roger S., Edzer Pebesma & Virgilio Gómez-Rubio. 2013. *Applied spatial data analysis with R* (Use R!). 2nd edn. New York: Springer.

Boberg, Charles. 2008. Regional phonetic differentiation in standard Canadian English. *Journal of English Linguistics* 36(2). 129–154. https://doi.org/10.1177/0075424208316648.

Boberg, Charles. 2019. A Closer Look at the Short Front Vowel Shift in Canada. *Journal of English Linguistics* 0075424219831353. https://doi.org/10.1177/0075424219831353.

Bowie, David. 2020. English in the North: The Vowels of Southcentral Alaska. In *Speech in the Western States: Volume 3, Understudied Varieties* (Publication of the American Dialect Society 105), 123–143. Durham, NC: Duke University Press. https://doi.org/10.1215/00031283-8820653 (5 December, 2020).

779   Cardoso, Amanda, Lauren Hall-Lew, Yova Kementchedjhieva & Ruaridh Purse. 2016. Between California and the Pacific

780   Northwest: The front lax vowels in San Francisco English. In Valerie Fridland, Tyler Kendall, Betsy E. Evans & Alicia

781   Beckford Wassink (eds.), *Speech in the Western States: Volume 1: The Coastal States* (Publication of the American

782   Dialect Society 101), 33–54. Durham, NC: Duke University Press. doi: 10.1215/00031283-3772890.

783   Carmony, Marvin. 1970. Some Phonological Rules of an Indiana Dialect. In Jerry Griffith & L. E. Miner (eds.), *The First*

784   *Lincolnland Conference on Dialectology*, 1–16. University, Alabama: University of Alabama Press.

785   Chang, Kang-Tsung. 2019. *Introduction to Geographic Information Systems*. 9th edn. New York: McGraw-Hill.

786   Clarke, Sandra, Ford Elms & Amani Youssef. 1995. The third dialect of English: Some Canadian evidence. *Language Variation*

787   *and Change* 7(2). 209–228. https://doi.org/10.1017/S0954394500000995.

788   Clayton, Ian & Valerie Fridland. 2020. Western Vowel Patterns in White and Native American Nevadans' Speech. In *Speech in*

789   *the Western States: Volume 3, Understudied Varieties* (Publication of the American Dialect Society 105), 39–63.

790   Durham, NC: Duke University Press. https://doi.org/10.1215/00031283-8820609.

791   Dodsworth, Robin & Mary Kohn. 2012. Urban rejection of the vernacular: The SVS undone. *Language Variation and Change*

792   24(2). 221–245. https://doi.org/10.1017/S0954394512000105.

793   Freeman, Valerie. 2014. Bag, beg, bagel: Prevelar raising and merger in Pacific Northwest English. *University of Washington*

794   *Working Papers in Linguistics* 32.

795   Grieve, J. 2013. A statistical comparison of regional phonetic and lexical variation in American English. *Literary and Linguistic*

796   *Computing* 28(1). 82–107. https://doi.org/10.1093/llc/fqs051.

797   Grieve, Jack. 2018. Spatial Statistics for Dialectology. In Charles Boberg, John Nerbonne & Dominic Watt (eds.), *The Handbook*

798   *of Dialectology*, 415–433. Hoboken, NJ, USA: John Wiley & Sons, Inc. https://doi.org/10.1002/9781118827628.ch24.

799   http://doi.wiley.com/10.1002/9781118827628.ch24 (23 July, 2019).

800   Gunter, Kaylynn, Ian Clayton & Valerie Fridland. 2017. Pre-velar Raising and Categorization in Nevada English. Presented at the

801   Third Annual Meeting of the NorthWest Phonetics & Phonology Conference, Vancouver, BC.

802   Hall-Lew, Lauren, Ruth Friskney & James M. Scobbie. 2017. Accommodation or political identity: Scottish members of the UK

803   Parliament. *Language Variation and Change* 29(3). 341–363. https://doi.org/10.1017/S0954394517000175.

804   Hinton, Leanne, Birch Moonwomon, Sue Bremner, Herb Luthin, Mary Van Clay, Jean Lerner & Hazel Corcoran. 1987. It's not

805   just the Valley Girls: A study of California English. In *Proceedings of the Thirteenth Annual Meeting of the Berkeley*

806   *Linguistics Society*, vol. 13, 117–128.

807   Jones, Jacqueline. 2015. *I Bag Your Pardon: The Albertan ae/ɛ Vowel Shift as a Window into Community Grammars*. University

808   of Calgary Thesis. http://dx.doi.org/10.5072/PRISM/26408. https://prism.ucalgary.ca/handle/11023/2717 (19 July,

809   2018).

810   Jones, Taylor. 2017. An application of geostatistics to AAE telephone survey data. Presented at the Annual Meeting of the American

811   Dialect Society, Austin, TX.

812   Kahle, David & Hadley Wickham. 2013. ggmap: Spatial Visualization with ggplot2. *The R Journal* 5(1). 144–161.

813   Kim, Chaeyoon, Sravana Reddy, James Stanford, Ezra Wyschogrod & Jack Grieve. 2019. Bring on the crowd! Using online audio

814   crowdsourcing for large-scale New England dialectology and acoustic sociophonetics. *American Speech* 94(2). 151--194.

815   https://doi.org/10.1215/00031283-7251252.

816   Krige, D. G. 1951. *A statistical approach to some mine valuations and allied problems at the Witwatersrand*. Johanesburg, South

817   Africa: University of Witwatersrand Master's Thesis.

818   Labov, William. 1963. The social motivation of a sound change. *WORD: Journal of the International Linguistic Association* 19(3).

819   273–309. https://doi.org/10.1080/00437956.1963.11659799.

820   Labov, William. 2006. *The social stratification of English in New York City*. 2nd edn. Cambridge: Cambridge University Press.

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

822   *change*. Berlin: Walter de Gruyter.

823   Labov, William, Malcah Yaeger & Richard Steiner. 1972. *A quantitative study of sound change in progress: Volume 1*.

824   Philadelphia, PA: US Regional Survey.

825   Leemann, Adrian, Marie-José Kolly & David Britain. 2018. The English Dialects App: The creation of a crowdsourced dialect

826   corpus. *Ampersand* 5. 1–17. https://doi.org/10.1016/j.amper.2017.11.001.

827   Leemann, Adrian, Marie-José Kolly, Jean-Philippe Goldman, Volker Dellwo, Ingrid Hove, Ibrahim Almajai, Sarah Grimm, Sylvain

828   Robert & Daniel Wanitsch. 2015. Voice Äpp: a mobile app for crowdsourcing Swiss German dialect data. In *Proceedings*

829   *of Interspeech*, 2804–2808. Dresden, Germany. https://doi.org/10.5167/uzh-114512.

830   Lenzo, Kevin. 2013. *The CMU Pronouncing Dictionary*. Carnegie Mellon University. http://www.speech.cs.cmu.edu/cgi-

831   bin/cmudict.

832   Manson, Steven, Jonathan Schroeder, David Van Riper & Steven Ruggles. 2018. *IPUMS National HIstorical Geographic*

833   *Information System*. [Database]. Minneapolis: University of Minnesota.

834   McLarty, Jason, Tyler Kendall & Charlie Farrington. 2016. Investigating the development of the contemporary Oregonian English

835   vowel system. In Valerie Fridland, Tyler Kendall, Betsy E. Evans & Alicia Beckford Wassink (eds.), *Speech in the*

836   *Western States, Vol. 1: The Pacific Coast* (Publication of the American Dialect Society 101), 135–157. Durham, NC:

837   Duke University Press. doi: 10.1215/00031283-3772934.

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

839   *Working Papers in Linguistics* 40(1). http://twpl.library.utoronto.ca/index.php/twpl/article/download/29162 (9 May,

840   2018).

Mielke, Jeff, Christopher Carignan & Erik R. Thomas. 2017. The articulatory dynamics of pre-velar and pre-nasal /æ/-raising in English: An ultrasound study. *The Journal of the Acoustical Society of America* 142(1). 332–349. https://doi.org/10.1121/1.4991348.

Nesbitt, Monica, Suzanne Wagner & Alexander Mason. 2019. A Tale of Two Shifts: Movement toward the Low-Back-Merger Shift in Lansing, Michigan. In Kara Becker (ed.), *The Low-Back-Merger Shift: Uniting the Canadian Vowel Shift, the California Vowel Shift, and short front vowel shifts across North America* (Publication of the American Dialect Society), vol. 104. Durham, NC: Duke University Press.

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

Purnell, Thomas C. 2008. Prevelar Raising and Phonetic Conditioning: Role of Labial and Anterior Tongue Gestures. *American Speech* 83(4). 373–402. https://doi.org/10.1215/00031283-2008-028.

R Core Team. 2018. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org.

Reed, Carroll E. 1952. The Pronunciation of English in the State of Washington. *American Speech* 27(3). 186–189. https://doi.org/10.2307/453476.

Reed, Carroll E. 1961. The Pronunciation of English in the Pacific Northwest. *Language* 37(4). 559–564. https://doi.org/10.2307/411357.

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

Rosen, Nicole & Crystal Skriver. 2015. Vowel patterning of Mormons in Southern Alberta, Canada. *Language & Communication* 42. 104–115. https://doi.org/10.1016/j.langcom.2014.12.007.

Stanley, Joseph. 2019. Phonological patterns in beg-raising. *UGA Working Papers in Linguistics* 4. 69--91.

Stanley, Joseph A. In press. A comparison of turn-of-the-century and turn-of-the-millennium speech in Georgia. *Proceedings of the 6th Annual Linguistics Conference at UGA*.

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).

Stanley, Joseph A. 2020. The Absence of a Religiolect among Latter-day Saints in Southwest Washington. In *Speech in the Western States: Volume 3, Understudied Varieties* (Publication of the American Dialect Society 105), 95–122. Durham, NC: Duke University Press. https://doi.org/10.1215/00031283-8820642 (5 December, 2020).

Swan, Julia Thomas. 2016. *Language Ideologies, Border Effects, and Dialectal Variation: Evidence from /æ/, /aʊ/, and /aɪ/ in Seattle, WA and Vancouver, B.C.* Chicago, Illinois: University of Chicago Ph.D. Dissertation.

872 Swan, Julia Thomas. 2020. Bag Across the Border. *American Speech* 95(1). 46–81. https://doi.org/10.1215/00031283-7587892.

873 Thiel, Anja & Aaron J. Dinkin. 2017. Escaping the TRAP: Losing the Nothern Cities Shift in Real Time. Presented at the New
874   Ways of Analyzing Variation 46, Madison, WI.

875 Thomas, Charles Kenneth. 1958. *An Introduction to the Phonetics of American English*. 2nd edn. New York: Ronald Press
876   Company.

877 Thomas, Erik R. 2003. Secrets Revealed by Southern Vowel Shifting. *American Speech* 78(2). 150–170.

878 Trudgill, Peter (ed.). 1984. *Applied Sociolinguistics* (Applied Language Studies). London: Academic Press, Inc.

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

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

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

886 Wassink, Alicia Beckford & Sharon Hargus. 2020. Heritage Language Features and the Yakama English Dialect. In *Speech in the
887   Western States: Volume 3, Understudied Varieties* (Publication of the American Dialect Society 105), 11–38. Durham,
888   NC: Duke University Press. https://doi.org/10.1215/00031283-8820598 (5 December, 2020).

889 Wassink, Alicia Beckford, Robert Squizzero, Mike Scanlon, Rachel Schirra & Jeff Conn. 2009. Effects of Style and Gender on
890   Fronting and Raising of /æ/, /e:/ and /ɛ/ before /g/ in Seattle English. Presented at the New Ways of Analyzing Variation
891   (NWAV) 38, Ottawa.

892 Wells, John C. 1982. *Accents of English: The British Isles*. Vol. 2. 3 vols. London: Cambridge University Press.

893 Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag. https://ggplot2.tidyverse.org.

894 Wickham, Hadley. 2017. *tidyverse: Easily Install and Load the "Tidyverse."* https://CRAN.R-project.org/package=tidyverse.

895 Yavaş, Mehmet. 2011. *Applied English Phonology*. Oxford, UK: Wiley-Blackwell. https://doi.org/10.1002/9781444392623.
896   http://doi.wiley.com/10.1002/9781444392623 (26 June, 2020).

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

899

900

901