HEARING *BEG* AND *BAG*: VARIATION IN THE PERCEPTION OF PRE-VELAR /Æ/-RAISING

Lisa Sullivan

University of Manitoba
lisa.sullivan@umanitoba.ca

ABSTRACT

Pre-velar /æ/-raising is a process in which some speakers of North American English raise /æ/ before /g/. Documentation of its production is limited, but even less is known about its perception. Perception studies using identification tasks [1, 2] found that perception does not vary based on listeners' production of the process. However, there is some evidence that individuals from regions without pre-velar /æ/-raising have greater metalinguistic awareness of it than those from regions with it [3], suggesting that there may be differences in perception. An AX discrimination task was conducted and revealed that individuals who raise less discriminated raised and unraised /æg/ more than those who raised more, suggesting that there is a relationship between perception and production of this process. Furthermore, differing results from previous studies suggest that this relationship is related to some aspect of perception that is captured by discrimination, but not identification, tasks.

Keywords: sociophonetics, sociophonetic perception, speech perception, North American English, pre-velar /æ/-raising

1. INTRODUCTION

Pre-velar /æ/-raising (hereafter, æg-raising) is a process in which some speakers of North American English raise (and front) the vowel /æ/ before the voiced velar stop (/g/). For speakers who æg-raise, the /æ/ in words like *bag* /bæg/ is pronounced higher than the /æ/ in words like *back* /bæk/, such that it may sound closer to the vowel in *beck* /bɛk/ or *bake* /bek/. In contrast, those who don't æg-raise would pronounce the /æ/ in *bag* very similarly to the vowel in *back*. Acoustic evidence shows that æg-raising occurs across Canada (e.g. [4, 5, 6, 7, 3]), as well as in the Pacific Northwest (e.g. [8, 9]), Upper Midwest (e.g. [10, 11]) and California [12, 13] in the US, but not in Nevada [14] or Colorado [3]. Outside these regions, evidence is limited to self-report data [15] which suggests that the prevalence of æg-raising decreases as you move away from the US-Canada border.

While documentation of the production of æg-raising is limited, even less is known about its perception. Freeman [1] and Sullivan [2] found no relationship between individual participants' production and perception of æg-raising in vowel identification tasks. However, evidence of greater metalinguistic awareness of æg-raising in non-æg-raising regions compared to æg-raising regions suggests that there may be a link that is not captured by these studies.

Variation in speech perception may be due to factors such as mobility, exposure and geography, as well as both the speaker's and listener's native dialects (e.g. [16, 17, 18]). The current study focuses on the link between production and perception. There is some evidence that there are links between perception and the listener's dialect. Listeners have been shown to classify sounds based on the phonemic system of their own dialect [19, 17, 20]. Furthermore, the perceptual saliency of certain dialect features is linked to listeners' native dialects [16, 19]. These dialect features are more salient to non-native dialect listeners than native ones. While these studies don't directly compare individual production and perception, differences between dialect regions suggests a possible link between production and perception.

Several studies [21, 22, 23] have shown that such a link may exist, even within dialects. Sumner and Samuel [23] found differences in the perception of /ɪ/-dropping between New York City speakers who /ɪ/-drop and those who don't. Fridland and Kendall [22] found that speakers' boundary between /e/ and /ɛ/ was correlated with their production, even within dialect. Evans and Iverson [21] found that Northern English speakers with more Northern English accents perceived vowels more closely to Northern productions than those who had more Southern productions. These studies suggest that differences in production within a dialect region may explain differences in perception within that region.

The current study uses a modified AX discrimination task to test if there is a relationship

between an individual's production of æg-raising and their ability to discriminate between raised and unraised /æg/. If individual production has an effect, participants who æg-raise less should differentiate raised and unraised /æg/ more than those who æg-raise more.

2. METHOD

2.1. Participants

49 participants, 25 from Ontario (13 F, 12 M) and 24 from Colorado (15 F, 9 M), completed the experiment. All participants were native speakers of English between the ages of 18 and 35 at the time of the study and who had spent at least 5 years of their childhood in their respective province/state. Ontario and Colorado were chosen as regions for this study as they were expected to produce a range of degrees of raising from no raising (Colorado) to large degrees of raising (Ontario) [24, 25].

2.2. Production Stimuli

The word list consisted of 60 monosyllabic English words including 22 words containing /æ/, /ɛ/ and /e/ followed by /g/ or /k/. See Sullivan [3] for a complete word list.

2.3. Perception Stimuli

Two 5-step (vowel quality) by 5-step (vowel duration) /æ/ to /ɛ/ continua were resynthesized in Praat [26] from the /ɛ/ natural tokens of a male native North American English speaker from the Midwest. The first continuum ranged from /bægsən/ to /bɛgsən/, and the second from /bæksən/ to /bɛksən/. These tokens served as X tokens for the AX discrimination task. The speaker's natural productions of /bægsən/ and /bæksən/ were used as A tokens. These tokens were run through the same script as the X tokens, except that the vowel quality and duration wasn't modified. Praat's [26] change gender feature was then used to make these sound female.

2.4. Procedure

Participants completed the tasks on their own computers using built-in or external microphones. They were asked to wear headphones for the perception task. They first completed the word list reading task in Gorilla [27]. Words were displayed on the screen individually, and participants said each word once. They completed the word list 3 times. The words were randomized each time.

Next, they complete the perception task in jsPsych [28]. For each trial, they heard the unraised A token of /bægsən/ or /bæksən/ followed by an X token from the corresponding continuum. They were asked to indicate if the second token was a different word from the first one, the same word as the first, but spoken with a different accent, or the same word as the first, spoken with the same accent. The 'same word, different accent' option was included so that participants could indicate when they were able to discriminate the two sounds, but they were not so different that they considered them to be different vowels (and thus different words) altogether. Participants heard each X token once (50 trials), and the order of items was randomized by participant.

3. ANALYSIS

3.1. Production

Production tokens were force aligned using the Montreal Forced Aligner [29] and manually corrected in Praat [26]. F1 and F2 measures were taken from the midpoint of each target vowel. Individual degree of raising for each speaker was calculated as a proportion of the F1 of /æk/ to the F1 of /ek/ using the formula in (1). Larger positive scores indicate more raising, as they indicate /æg/ is closer to /ek/. Scores close to 0 indicate no raising, and negative scores indicate that /æg/ is lower than /æk/.

$$(1) \qquad \frac{M_{\text{æk}} - M_{\text{æg}}}{M_{\text{æk}} - M_{\text{ek}}}$$

3.2. Perception

Since the primary research question was about whether or not participants were able to distinguish /æg/ from /æk/, participant responses to the discrimination class were collapsed into two categories "Same", which includes the response "Same word with the same accent" and "Different" which includes both "Different word" and "Same word with a different accent" for the purposes of the statistical analysis. However, the 3 way response variable is considered descriptively, as well.

A mixed effects logistic regression model was constructed with discrimination as the response variable using the *glmer*() function from the lme4 [30] package in R [31]. The lmerTest package [32] was used to obtain significance values. Predictor variables, interactions and random effects were

included as listed in (2).

(2) glmer(Discrimination ~(Vowel Quality + Duration) * Phonological Context * (Production + Metalinguistic Awareness) + (1 + Phonological Context + Vowel Quality + Duration | Participant), data = Perception Data, family = "binomial", control = glmerControl(optimizer = "bobyqa", optCtrl=list(maxfun=2e6)))

The continuous variables (duration, vowel quality, metalinguistic awareness and production) were z-score normalized. Phonological context coded with /k/ as -0.5 and /g/ as 0.5. For post-hoc analysis and descriptive statistics, participants divided into low and high æg-raising groups based on the mean degree of æg-raising.

4. RESULTS

4.1. Production

Individual degree of raising scores are shown in Figure 1. While Ontarians raise more than Coloradans, there is a large degree of variation both within and across groups, ranging from no raising, to raising /æg/ over 75% of the way from /æk/ to /ek/.

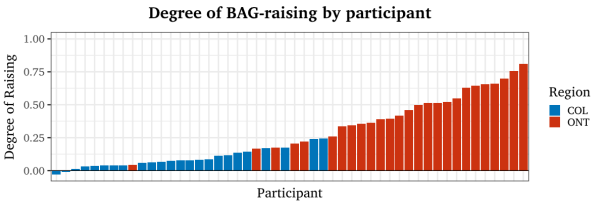

Figure 1: Individual degree of BAG-raising

4.2. Perception: 2-way response variable

Overall, the results are consistent with what would be found if participants were doing the task as expected. X tokens are discriminated from A tokens less when the vowel quality is more /æ/-like (Est. = -1.7, *SE* = -.11, z = -13.48, p < 0.001) and duration is longer (Est. = -1.4, *SE* = -.10, z = -13.59, p < .001).

There are overall differences in how the vowel is distinguished before /g/ and /k/ in line with the fact that the distinction in production only happens before /g/. Participants are more likely to distinguish more /ɛ/-like tokens from unraised /æ/ when the vowels occur before /k/, where the process doesn't happen, than before /g/, where it does. This is born out in the interaction between vowel quality and phonological context in the regression model (Est. = 1.58, *SE* = .15, z = 10.30, p < .001).

This effect is modulated by individual production. As shown in Figure 2, participants who æg-raise less show similar levels of discrimination of more /ɛ/-like tokens before /g/ and /k/. On the other hand, those who æg-raise more show a much lower level of discrimination before /g/, than /k/, suggesting they are less sensitive to distinctions between /ɛg/ and /æg/. This is born out in a three way interaction between vowel quality, phonological context and production in the regression model (Est. = 0.59, *SE* = .15, z = 3.83, p < .001). Consistent with what Figure 2 shows, post-hoc analyses shows a larger effect size of the interaction between vowel quality and phonological context for the high æg-raising group (Est. = 2.09) than the low one (Est. = 1.20).

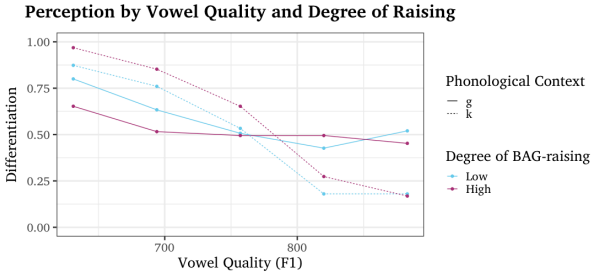

Figure 2: Discrimination Results (2 Levels)

4.3. Perception: 3-way response variable

For the statistical analysis, the three way response variable was collapsed into two levels; however, in 29.5% of the "different" responses participants indicated that the vowel was the same as in the natural unraised /æ/ token, but spoken in a different accent, rather than being a different vowel altogether. Figure 3 shows the proportion of participant responses divided by phonological context, low or high degree of BAG-raising in perception and vowel quality of the vowel in the continuum token.

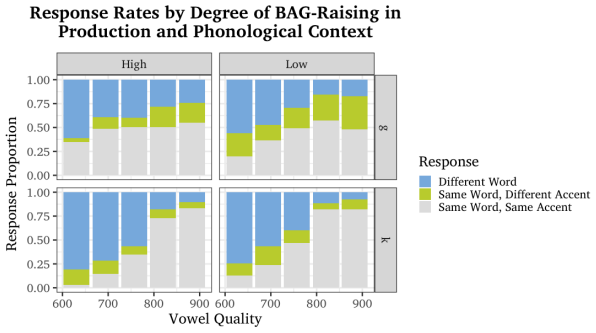

Figure 3: Discrimination Results (3 Levels)

The different accent responses in the /k/ context are similar for both groups, and occur at a lower rate

than in the /g/ context, where they seem to occur more with more /æ/-like tokens, perhaps because speakers are less willing to recategorize these tokens than the more /ɛ/-like ones. Furthermore, low æg-raisers were more likely to have different accent responses in the /g/ context, especially for the more /ɛ/-like tokens, suggesting that these speakers differentiate more in this context, even if they don't recategorize.

In summary, the results show that (1) participants are more likely to differentiate a vowel from /æ/ the less /æ/-like it is, (2) they are more likely to discriminate less /æ/-like vowels before /k/ than /g/, (3) high æg-raisers discriminate less before /g/ than low æg-raisers.

5. DISCUSSION

The goal of this study was to determine if there is a relationship between an individual's production of æg-raising and their ability to discriminate between raised and unraised /æg/. The results indicate that participants are able to differentiate variants of /æg/ and that this is modulated by individual production of æg-raising. Furthermore, the use of a discrimination task with a three way response variable may have contributed to detecting this difference.

The results of this study are in line with the metalinguistic evidence [3] that individuals from non-æg-raising regions are more aware of æg-raising than those from æg-raising regions, as both findings show greater sensitivity to æg-raising for those who do not æg-raise themselves or aren't exposed to it in the community. Furthermore, the results are consistent with previous findings on other phenomena that show a link between production (or dialect) and perception, particularly those that show links between individuals' production and perception [21, 22, 23] and those that show non-native dialect speakers have greater perceptual saliency of certain dialect features than native dialect speakers [16, 19].

On the other hand, the results run contrary to previous æg-raising studies which found no link between production and perception [1, 2]. There are a number of methodological considerations which could explain the difference between this study and the previous ones, including a higher number of participants, the possibility that these participants had a larger range of degrees of raising, and differences in the synthesis and choice of stimuli; however, the most likely explanation seems to be in the task itself.

Where as Freeman [1] and Sullivan [2] used identification tasks which required the participant to categorize the vowel they were perceiving as an English vowel, the current task used a discrimination task and allowed participants to note within-category differences. It could be that the nature of the tasks (discrimination vs identification), explains why the results are different. Perhaps, discrimination better taps into the aspect of perception that results in differences in metalinguistic awareness.

However, it is possible that the third option in the task, which allowed the participant to say that the second word was the same as the first, but just had a different accent, allowed for this difference to emerge. This seems to be supported by the 3-way categorization results where much of the difference between the low and high æg-raising groups in the /g/ context arises from the different accent responses, and not the different word responses. It may, therefore, be that the link between perception and production doesn't have to do with recategorization of the vowel, but rather with hearing it as a variant of the same vowel, a distinction the discrimination tasks previously employed weren't able to capture.

6. CONCLUSION

The current study finds that there appears to be a link between the production and perception of æg-raising. It also suggests that perception studies of dialectal variation should incorporate options for identifying variants of the same vowel, and not just perception of different vowels. Future work should continue to examine how æg-raising is perceived, including examining its perception in context, and with real words. Furthermore, an identification task which allows for identification of variants of the same vowel could be developed to tease apart possible reasons for the difference in results between this study and previous ones [1, 2].

7. REFERENCES

[1] V. Freeman, "Prevelar merger in production vs. perception," in *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019*, S. Calhoun, P. Escudero, M. Tabain, and P. Warren, Eds. Canberra, Australia: Australasian Speech Science and Technology Association Inc., 2019, pp. 457–461.

[2] L. Sullivan, "The production and perception of prevelar/æ/-raising by Canadian and American English speakers," *The Journal of the Acoustical Society of America*, vol. 148, no. 4, p. 2809, 2020. [Online]. Available: https://doi.org/10.1121/

1.5147827
[3] ——, "Pre-velar /æ/-raising in Ontario and Colorado English: Production, perception and metalinguistic awareness," Ph.D. dissertation, Univesrity of Toronto, Toronto, ON, 2022.
[4] C. Boberg, "Regional phonetic differentiation in standard Canadian English," *Journal of English Linguistics*, vol. 36, no. 2, pp. 129–154, 2008.
[5] J. H. Esling and H. J. Warkentyne, "Retracting of/æ/in Vancouver English," in *Focus on Canada*, S. Clarke, Ed. John Benjamins, 1993, vol. 11, pp. 229–246.
[6] S. Onosson, "Prevelar vowel raising and merger in Manitoba English," *Journal of English Linguistics*, vol. 50, no. 3, pp. 247–280, 2022.
[7] N. Rosen and C. Skriver, "Vowel patterning of mormons in southern Alberta, Canada," *Language & Communication*, vol. 42, pp. 104–115, 2015.
[8] V. Freeman, "Vague eggs and tags: Prevelar merger in Seattle," *Language Variation and Change*, vol. 33, no. 1, pp. 57–80, 2021.
[9] J. M. Riebold, "The social distribution of a regional change:/æg, ɛg, eg/in Washington State," Ph.D. dissertation, University of Washington, 2015.
[10] E. Koffi, "The acoustic vowel space of central Minnesota English: Focus on female vowels," *Linguistic Portfolios*, vol. 2, no. 1, pp. 2–16, 2013.
[11] C. Zeller, "The investigation of a sound change in progress: /ae/to/e/in midwestern American English," *Journal of English linguistics*, vol. 25, no. 2, pp. 142–155, 1997.
[12] A. Cardoso, L. Hall-Lew, Y. Kementchedjhieva, and R. Purse, "Between California and the Pacific Northwest: The front lax vowels in San Francisco English," *American Speech: A Quarterly of Linguistic Usage*, vol. 2016, pp. 33–54, 2016.
[13] L. Esposito and E. Lake, "Complicating prevelar raising in the West," *American Speech*, pp. 1–45, 07 2021.
[14] K. Gunter, I. Clayton, and V. Fridland, "Pre-velar raising and categorization in Nevada," in *New Ways of Analyzing Variation 47*, New York City, N.Y., 2018.
[15] J. A. Stanley, "Regional patterns in prevelar raising," *American Speech*, vol. 97, no. 3, pp. 374–411, 2022.
[16] R. Allbritten, "Sounding Southern: Phonetic features and dialect perceptions," Ph.D. dissertation, Georgetown University, Washington, DC, 2011.
[17] B. O. Flanigan and F. P. Norris, "Cross-dialectal comprehension as evidence for boundary mapping: Perceptions of the speech of southeastern Ohio," *Language Variation and Change*, vol. 12, no. 2, pp. 175–201, 2000.
[18] J. Hay and K. Drager, "Stuffed toys and speech perception," *Linguistics*, vol. 48, no. 4, pp. 865–892, 2010.
[19] C. G. Clopper, J. B. Pierrehumbert, and T. N. Tamati, "Lexical neighborhoods and phonological confusability in cross-dialect word recognition in noise," *Laboratory Phonology*, vol. 1, no. 1, pp. 65–92, 2010.
[20] E. Jacewicz and R. A. Fox, "The effects of cross-generational and cross-dialectal variation on vowel identification and classification," *The Journal of the Acoustical Society of America*, vol. 131, no. 2, pp. 1413–1433, 2012.
[21] B. G. Evans and P. Iverson, "Plasticity in vowel perception and production: A study of accent change in young adults," *The Journal of the Acoustical Society of America*, vol. 121, no. 6, pp. 3814–3826, 2007.
[22] V. Fridland and T. Kendall, "Exploring the relationship between production and perception in the mid front vowels of US English," *Lingua*, vol. 122, no. 7, pp. 779–793, 2012.
[23] M. Sumner and A. G. Samuel, "The effect of experience on the perception and representation of dialect variants," *Journal of Memory and Language*, vol. 60, no. 4, pp. 487–501, 2009.
[24] C. Boberg, *The English language in Canada: Status, history and comparative analysis*. Cambridge University Press, 2010.
[25] C. Holland and T. Brandenburg, "Beyond the front range: The Coloradan vowel space," *Publication of the American Dialect Society*, vol. 102, no. 1, pp. 9–30, 2017.
[26] P. Boersma and D. Weenink, "Praat: doing phonetics by computer [computer program]. version 6.1.38," 2021. [Online]. Available: http://www.praat.org/
[27] A. L. Anwyl-Irvine, J. Massonnié, A. Flitton, N. Kirkham, and J. K. Evershed, "Gorilla in our midst: An online behavioral experiment builder," *Behavior research methods*, vol. 52, no. 1, pp. 388–407, 2020.
[28] J. R. de Leeuw, "jspsych: A javascript library for creating behavioral experiments in a web browser," *Behavior Research Methods*, vol. 47, no. 1, pp. 1–12, 2015.
[29] M. McAuliffe, M. Socolof, S. Mihuc, M. Wagner, and M. Sonderegger, "Montreal Forced Aligner: Trainable text-speech alignment using Kaldi." in *Interspeech*, vol. 2017, 2017, pp. 498–502.
[30] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.
[31] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2020. [Online]. Available: https://R-project.org/
[32] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "lmerTest package: Tests in linear mixed effects models," *Journal of Statistical Software*, vol. 82, no. 13, pp. 1–26, 2017.

Acknowledgements: This work draws on research supported by the Social Sciences and Humanities Research Council of Canada.