(CANADIAN) RAISE YOUR VOWELS IN SONG

Murray Schellenberg* and **Avery Ozburn†**

Department of Linguistics, University of British Columbia

1   Introduction

One of the iconic features of Canadian English is the raising of the diphthongs [aj] and [aw] to [ʌj] and [ʌw] before voiceless obstruents [1]. This process, known as Canadian raising (CR), is phonological and may occur regardless of speech rate [2], and it has been claimed that it arises from the shortening effect these consonants have on preceding vowels [3]. It is currently unknown to what extent CR occurs in singing, in which large differences in note duration make it impossible to correct for speech rate. This question is particularly interesting for professional singers, who are trained to modify their vowels. Trained singers have been shown to have a significantly reduced vowel space [4] and to maintain a more open jaw posture [5], which may interfere with raising. In this paper, we investigate to what extent CR exists in the singing of trained Canadian singers, and whether and how it is different from their raising in normal speech. We report on an experiment in which Canadian singers are asked to say and sing passages containing multiple tokens of raising vowels followed by voiceless and voiced consonants. Tokens are embedded in novel lyrics written to fit commonly known melodies. We test the degree of raising in singing compared to speaking as well as within each of these production modes. Results contribute to our understanding of the acoustics of singing and how singer dialect interacts with the acoustic realization of the articulatory settings of trained singing.

2   Methods

To test CR in singing, we wrote new lyrics to two familiar melodies, in order to include a large number of raising diphthongs ([aj] and [aw]) in raising contexts (followed by voiceless obstruents), as well as in non-raising contexts (followed by voiced obstruents) for comparison. In the lyrics, we had 31 tokens of [aj] followed by a voiceless obstruent, 16 of [aj] followed by a voiced obstruent, 15 of [aw] followed by a voiceless obstruent, and 8 of [aw] followed by a voiced obstruent.

We recruited 3 professionally trained singers (1F/2M) who are native speakers of Canadian English. They were seated in a quiet room, in front of a Macintosh desktop computer and Yamaha YST-M40 speakers, and they wore a C520 headset microphone plugged into an M-Audio Fast Track Pro pre-amp. Sound was recorded into Audacity.

The experiment was presented using PowerPoint. The two songs were presented one at a time, first for speaking, then listening to the music, then singing.

In the speaking section, the lyrics were written out on a slide, with one verse per slide. Instructions told participants to say the text aloud three times before moving to the next slide. Following all of the verses, the music slide came up, with embedded sound that participants could click to listen to the music. They were permitted to do so as many times as they wished. When ready, they proceeded to the singing slides. On each slide, the lyrics to one verse were given, together with an embedded sound icon for the music for that verse. They could listen to the music as many times as they wished, and were instructed to sing each verse three times before moving on to the next slide. Note that because the headset microphone was directional, the music playing from the speakers could not be heard in the recording. Following the completion of the first song, the same procedure applied to the second. After both songs, participants completed a short background questionnaire.

3   Results

All diphthongs of interest were segmented manually in Praat [6], and a script was used to extract F1 and F2 values at the 25% point. These values were then used to determine subjects' raising.

We excluded the data from two of the participants (both males) for not raising in the spoken sections. The lack of raising was apparent both auditorily (to trained phoneticians who are native speakers of Canadian English) and through the formants, where F1 was slightly though not significantly higher in raising contexts than in non-raising contexts.[1] The remaining discussion focuses on the remaining (female) participant.

For all speaking/singing and raising/non-raising conditions, the difference in mean F1 between the two vowels [aj] and [aw] (at the 25% point that we measured) were quite small (the largest was less than 40Hz). We concluded that the initial vowel of both diphthongs is essentially equivalent in height for this particular speaker and consequently combined the vowels [aj] and [aw] in further analysis, to get a general picture of this participant's raising. The mean F1 and F2 values for this speaker in each of the contexts and conditions are given in Table 1.

Table 1: Formant values in each condition

Mode of Speech	Raising Context?	F1 (Hz)	F2 (Hz)
Spoken	Yes	566.03	1323.67
Spoken	No	620.44	1306.28
Sung	Yes	518.32	1272.11
Sung	No	563.65	1296.50

* mhschellenberg@gmail.com
† aozburn@gmail.com

As is evident, raising is primarily seen in F1, with F1 values lower for the raising contexts than the non-raising contexts. This is true for both the spoken condition and the sung condition. Further, the F1 values for singing are lower than those of speech for both raising and non-raising contexts.

Despite having only one speaker, we ran a linear model to determine whether any of these effects neared significance, and to look at whether singing and speech have similar degrees of raising. The dependent variable was F1, and the independent variables were mode of speech (sung versus spoken), following consonant voicing (raising versus non-raising context), and their interaction. The reference condition was the spoken mode in the non-raising context.

The factor of raising neared significance (p = 0.0623), as did the factor of mode of speech (p = 0.0918). We expect that these effects would reach significance with more speakers. In both cases, the estimates and t-values were negative, indicating that raising and singing both (independently) lower F1. The interaction effect did not near significance (p = 0.8265). The estimate and t-value for the interaction was positive, indicating a very slight, far-from-significant trend for less raising in singing than in speech.

4 Discussion and Conclusions

The results show substantial F1 differences between raising and non-raising contexts in the singing condition (518.32Hz vs. 563.65Hz), suggesting that this singer maintains CR during singing. This singer has 25 years of training and experience singing and has sung professionally for seven years. Her maintenance of CR in spite of her training speaks to the robustness of this pattern even in singing. In particular, training for singers includes training on English diction. LaBouff [7] is one of the standard texts for this type of training. She instructs singers to sing in "neutral English", which she defines (p. 6) as one of three accents: American Standard, RP, or Mid-Atlantic. (The choice depends on repertoire, with British music in RP, North American music in GA, and anything else in Mid-Atlantic.) The singer reported training of this sort, with a strong emphasis on singing English with a Mid-Atlantic dialect. However, despite this training, CR persists in her singing

Further, raising is maintained to an approximately equivalent degree in singing as in speaking. However, there is a slight (non-significant) decrease in raising magnitude, with the F1 mean in raising sung contexts about 9 Hz higher than we would expect from a simple combination of the F1 decrease from raising alone combined with that from singing alone. We were only able to look at the results of a single singer; further research is needed to investigate how robust this difference is, and whether it becomes significant with additional singers.

For the preliminary analysis, we proceeded with an assumption that [aj] and [aw] start at the same vowel, and we pooled the results across diphthongs. While means suggested this assumption is fairly accurate for this speaker, it may not hold more generally [8], and so future work should look at the diphthongs separately.

This preliminary study did not consider duration, which may also contribute to the presence of CR in singing. Diphthongs set to longer notes are traditionally sung on the initial vowel of the diphthong in Western classical music, which may mean that longer notes show a lesser degree of CR. Preliminary auditory impressions suggest that longer diphthongs may, in fact, have less CR than those set to shorter notes. However, we do not presently have enough data to confirm this impression.

There was an overall lowering of F1 in singing compared to speaking, which neared significance even for the single speaker. As a result, the F1 values for the singer's unraised sung vowels are surprisingly similar to those of her raised spoken vowel values. This F1 lowering may well be a result of the smaller vowel space that results from Western classical singing technique [4]: the singer may be lowering her whole vowel space to compensate for the lowering of higher vowels that is common in singing.

Overall, this study has shown that CR is maintained in sung speech by at least one professional singer who is a native speaker of Canadian English. The results are suggestive of the persistence of singer dialect despite training, and raise a number of possibilities for future research into this phenomenon.

Endnotes

[1.] One of the speakers, though born and raised in Canada, had British parents and a slight British accent. We have no explanation for the lack of raising in the other speaker.

Acknowledgments

References

[1] J.C.Wells *Accents of English* (Vol. 3). Cambridge University Press, 1982.

[2] P. Boersma & J. Pater. Constructing constraints from language data: The case of Canadian English diphthongs. *Handout, NELS*, *38*, 2007.

[3] J. Myers. Canadian raising and the representation of gradient timing relations. *Studies in the Linguistic Sciences,* 27(1), 1997.

[4] W. Ophaug. *Sangfonetikk: en innføring*. Fagbokforlaget, 2010.

[5] A. Nair, G. Nair & G. Reishofer. The Low Mandible Maneuver and Its Resonential Implications for Elite Singers. *Journal of Voice*, *30*(1), 2016.

[6] P. Boersma, and D. Weenink. Praat: doing phonetics by computer [Computer program]. Version 5.4.08, http://www.praat.org/, 2015.

[7] K. LaBouff, *Singing and communicating in English: a singer's guide to English diction*. Oxford University Press. 2007.

[8] R. E. Hagiwara. Vowel production in Winnipeg. *The Canadian Journal of Linguistics/La revue canadienne de linguistique*, *51*(2), 127-141, 2006.