The effect of pitch, intensity, and vowel quality on lip aperture in the singing voice

It is a known fact in phonetic studies that loudness (or intensity) and pitch (or fundamental frequency) are typically interdependent (Sundberg, 1987). A telling example is the act of screaming, where an individual will tend to increase their fundamental frequency as they increase the loudness of their scream, and this in turn may necessitate a bigger buccal aperture (Schwartz et al., 2020). The aim of this work is to study the influence of loudness and pitch on mouth opening in another specific vocal use: singing. From a practical or musicological perspective to the scientific study of singing voice processes, most of the literature agrees that there is an undeniable link between pitch and intensity (Kayes, 2019), which tends to impact buccal aperture as well (Cornut 2002, Scotto di Carlo, 2005). Though it is an essential parameter to take into account when singing, no study has actually attempted to test the effect of loudness and pitch on mouth opening, which we propose to do here. Additionally, we looked at the role of vowel quality and if any effect of close vs. open vowels on mouth aperture (as those found in speaking) could be identified in singing. Considering the limited data available, this approach remains at a more descriptive and exploratory stage.

Three professional male British contemporary singers were recorded singing a total of 21 10-second sentences created for the purpose of the experiment. The sentences were divided up into 3 different 80's rock-inspired melodies. The 21 sentences were sung twice, once at a high pitch and once at a lower pitch. The pitches were individually determined beforehand, on the basis of the highest note the singer could produce comfortably – this note was selected to be the highest possible one in the three melodies for the high pitch category, and was transposed one octave below for the low pitch category. They were also filmed. Fundamental frequency (Hz) and intensity (dB) data was extracted through Praat (Boersma and Weenink, 2024). For mouth opening, a post-hoc AI-based face landmark detection video treatment (Google AI, 2024) was applied (see Figure 1), and the Euclidean distance in pixels between the upper-most middle point of the upper lip and the lower-most middle point of the lower lip was calculated and converted to cm. A linear mixed-effects regression model was fitted in R, with lip opening as the dependent variable, and fundamental frequency and intensity as predictors, and their interaction. Random intercepts were included for sentence tokens and participants. Both fundamental frequency and intensity showed significant positive effects ($\beta = 0.301$, $SE = 0.036$, $t = 8.403$, $p < 0.001$ and $\beta = 0.139$, $SE = 0.017$, $t = 8.147$, $p < 0.001$ respectively) on buccal aperture. The interaction between the two factors was also statistically significant. Figure 2 shows the partial dependency of the results as a heatmap: the lighter the color, the higher the predicted lip aperture based on the combination of pitch and intensity. As such, the predicted effect of fundamental frequency and intensity on mouth aperture appears to be corroborated by the data. For the second step of the analysis, looking more closely at vowel quality, 2 × 5 words with open vowels /ɑː/ and /æ/ and 2 × 5 words with close vowels /iː/ and /ɪ/ were identified in the lyrics, the vowels segmented in Praat and their corresponding fundamental frequency, intensity and mouth aperture measures extracted. Figure 3 shows that there seems to be a greater possible range of mouth aperture for open vowels as opposed to close vowels, and the data for s03 quite tellingly reveals a strong difference in average aperture.

These findings serve as a basis for further research on articulation and pronunciation in the singing voice, where certain sounds may be favored over others in certain conditions, such as in higher-pitched passages.

Figure 1. Face landmark detection

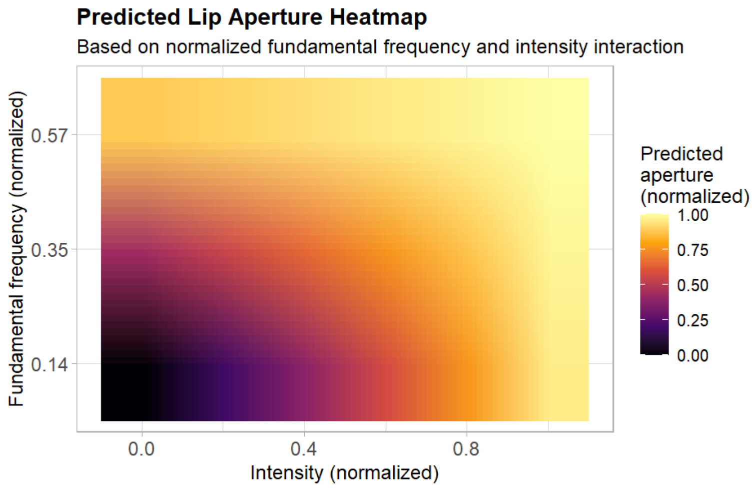

Figure 2. Partial dependency plot

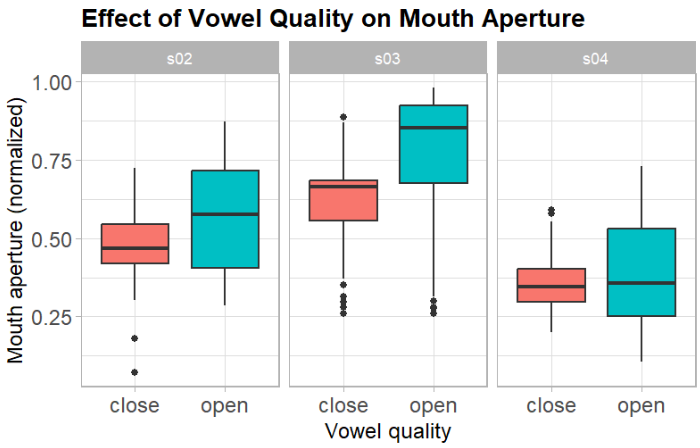

Figure 3. Mouth aperture according to vowel quality per participant

Keywords singing voice, fundamental frequency, intensity, vowel quality

References

BOERSMA, P., & WEENINK, D. (2024). *Praat: Doing phonetics by computer* (Version 6.3) [Computer software]. Retrieved from http://www.praat.org

CORNUT, G. (2002). *Moyens d'investigation et pédagogie de la voix chantée*. Symétrie.

GOOGLE. (2004). *Face Landmarker*. Google AI. Retrieved January 2, 2025, from https://ai.google.dev/edge/mediapipe/solutions/vision/face_landmarker

KAYES, G. (2019). Structure and function of the singing voice in Welch, G., Howard, D. and Nix, J. (eds), *The Oxford Handbook of Singing* (pp. 1–42). *Oxford University Press*.

R CORE TEAM. (2024). *R: A language and environment for statistical computing* (Version 4.3) [Computer software]. R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

SCOTTO DI CARLO, N. (2005). Contraintes de production et intelligibilité de la voix chantée. *Travaux Interdisciplinaires du Laboratoire Parole et Langage d'Aix-en-Provence (TIPA)*, *24*, 159–180.

SUNDBERG, J. (1987). *The Science of the Singing Voice*. Northern Illinois University Press.

SUNDBERG, J. (1990). What's so special about singers? *Journal of Voice*, *4*(2), 107–119.

SCHWARTZ, J. W., ENGELBERG, J. W. M., & GOUZOULES, H. (2020). Was That a Scream? Listener Agreement and Major Distinguishing Acoustic Features. *Journal of Nonverbal Behavior*, *44*(2), 233–252.