Vowel Content Influences Relative Pitch Perception in Vocal Melodies

Frank A. Russo
Ryerson University, Toronto, Canada

Dominique T. Vuvan
Skidmore College

William Forde Thompson
Macquarie University, Sydney, Australia

Note-to-note changes in brightness are able to influence the perception of interval size. Changes that are congruent with pitch tend to expand interval size, whereas changes that are incongruent tend to contract. In the case of singing, brightness of notes can vary as a function of vowel content. In the present study, we investigated whether note-to-note changes in brightness arising from vowel content influence perception of relative pitch. In Experiment 1, three-note sequences were synthesized so that they varied with regard to the brightness of vowels from note to note. As expected, brightness influenced judgments of interval size. Changes in brightness that were congruent with changes in pitch led to an expansion of perceived interval size. A follow-up experiment confirmed that the results of Experiment 1 were not due to pitch distortions. In Experiment 2, the final note of three-note sequences was removed, and participants were asked to make speeded judgments of the pitch contour. An analysis of response times revealed that brightness of vowels influenced contour judgments. Changes in brightness that were congruent with changes in pitch led to faster response times than did incongruent changes. These findings show that the brightness of vowels yields an extra-pitch influence on the perception of relative pitch in song.

Received: May 7, 2018, accepted May 30, 2019.

Key words: singing, vowels, timbre, relative pitch, psychoacoustics

A number of factors extraneous to fundamental frequency have influence over the perception of relative pitch. For example, the perception of relative pitch can be influenced by pitch register (Russo & Thompson, 2005a), loudness (McDermott, Lehr, & Oxenham, 2008; Thompson, Peter, Olsen, & Stevens, 2012), facial movement (Abel, Li, Russo, Schlaug, & Loui, 2016; Thompson, Graham, & Russo, 2005; Thompson, Russo, & Livingstone, 2010), and brightness (Allen & Oxenham, 2014; McDermott et al., 2008; Russo & Thompson, 2005b). With regard to the latter, note-to-note changes in brightness that are congruent with changes in pitch have been found to lead to an expansion of perceived interval size (i.e., pitch distance), whereas incongruent changes lead to a contraction.

The brightness of a sung note can be dynamically manipulated through performance expression with effects on audibility (Sundberg, 1972; 1994) as well as emotion (Livingstone, Choi, & Russo, 2014). The brightness of a sung note can also be manipulated through word selection, and through vowel content in particular. In the case of song, pitch is produced in a manner that is determined by the melody and should thus be independent of vowel content; however, on the basis of other research demonstrating extra-pitch influences, it seems likely that vowel content has some influence over relative pitch perception. In the current study, we investigate whether note-to-note changes in brightness arising from vowel content may influence the perception of relative pitch.

Although the musical significance of vowel content has been explored with regard to timbral aspects of music (e.g., Slawson, 1985), to our knowledge, only one prior study has done so with regard to the perception of pitch relations. Fowler and Brown (1997) investigated the pitch separation that is necessary to hear a "high vowel" as being equal in pitch to a "low vowel." Using naturally produced spoken vowels that were resynthesized to manipulate fundamental frequency, Fowler and Brown found that a high vowel [i] had to exceed a low vowel [a] by 4.35 Hz in order to sound equal in pitch. This finding maps on to the notion that [i] sounds "higher" than [a]. In phonetic terms, "high vowel" refers to the height of the tongue relative to the roof of the mouth. In acoustic terms, the range of formants in high vowels tend to be relatively wide but the first formants (F1) tend to be relatively low (Hillenbrand, Getty, Clark, & Wheeler, 1995). The net result of these formant differences is that high

Music Perception, volume 37, issue 1, pp. 57–65, issn 0730-7829, electronic issn 1533-8312. DOI: https://doi.org/10.1525/mp.2019.37.1.57

vowels tend to have reduced brightness relative to low vowels.

Although brightness has been associated with several spectral features, spectral centroid appears to be the most robustly linked acoustic dimension (Hall & Beauchamp, 2009; McAdams & Giordano, 2011; Schubert & Wolfe, 2006). In order to characterize relative brightness for stimuli with varying pitch height, we defined our stimuli in regards to the normalized spectral centroid, quantified as the amplitude-weighted mean of the frequency spectrum divided by fundamental frequency.

Experiment 1a

In Experiment 1a, we investigated whether note-to-note changes in brightness arising from vowel content can influence the perception of sung interval size. Three-note sequences were synthesized using vocal synthesis, allowing for independent control over brightness and pitch. Each vowel was produced in the context of a consonant-vowel syllable: d[i], d[o̜], d[ɑ]. See Figure 1 for examples of vowel spectra, along with associated spectral centroids (f_c).

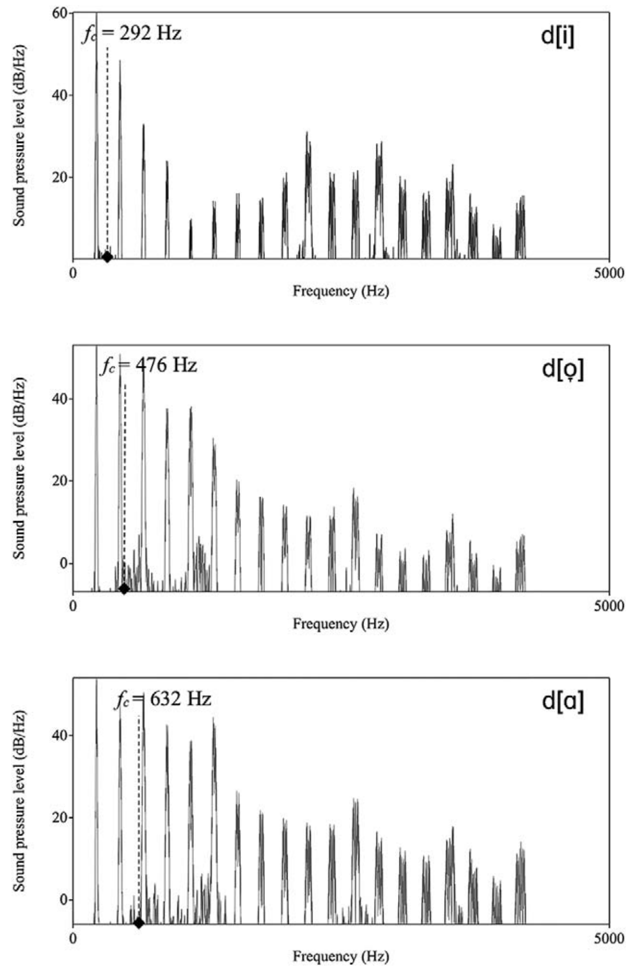

FIGURE 1. Vowel spectra and associated spectral centroids for syllables (di, do, da) synthesized on A_3 (220 Hz).

Our key prediction was that pitch and brightness would interact, such that trials in which pitch and brightness changed in the same direction (congruent trials) would elicit the perception of a larger interval than trials in which pitch and brightness changed in opposite directions (incongruent trials). Furthermore, we predicted that larger changes in brightness across the sequences would lead to a larger congruency effect than smaller changes in brightness.

Method

PARTICIPANTS

Twenty-seven participants ranging in formal music training from 0 to 9 years ($M = 2.69$ years of formal training, $SD = 2.91$), were recruited from an introductory psychology course at the University of Toronto, Mississauga. These participants included 11 men and 16 women, ranging in age from 17 to 28 ($M = 19.44$ years, $SD = 2.29$). All reported having normal hearing and received course credit for their participation.

STIMULI

To minimize uncontrolled variability, note sequences were synthesized rather than produced by real vocalists. Synthesis was realized using VocalWriter 2.0 software (KAE Labs, 2005), which is based on the Klatt formant synthesizer (Klatt & Klatt, 1990). Three consonant-vowel syllables (d[i], d[o̜], d[ɑ]; hereon: "di," "do," "da") were synthesized on all chromatic pitches falling between F_3 and E_4. Each syllable was 1.22 s in total duration. The consonant portion of the syllable ("d") was 20 ms and the offset portion of all vowels was to 7 ms. Figure 1 provides examples of vowel spectra and associated centroids for syllables synthesized on A_3 (220 Hz). The frequency normalized spectral centroid (i.e., ratio of the spectral centroid to the fundamental frequency) for all pitches used in Experiment 1 was relatively low for di, high for da, and intermediate for do (see Figure 2).

Syllable sandwiches. Syllables were combined to form "syllable sandwiches." In each case, two identical syllables produced at the same pitch surrounded a central syllable produced at different pitch (da-di-da, di-da-di, da-do-da, do-da-do, do-di-do, and di-do-di). The

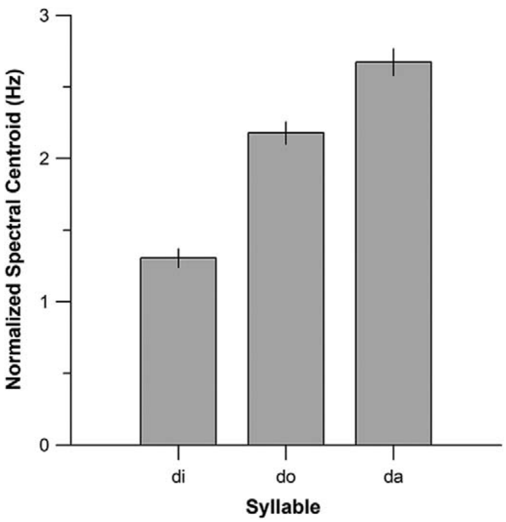

FIGURE 2. Spectral centroid (with standard error bars) of di, do, and da (normalized for f_0)

change in brightness was largest for syllable sandwiches combining da and di, smallest for syllable sandwiches combining da and do, and intermediate for syllable sandwiches combining do and di. This factor is referred to as *brightness change*. The sequencing of syllables led to a rise-fall or fall-rise contour. This factor is referred to as *brightness contour*. The pitch change between the first and second tone was equivalent to the pitch change between the second and third note, corresponding to a perfect fifth (P5, seven semitones) or a tritone (TT, six semitones). This factor is referred to as *pitch change*. The sequence of pitches across the three notes led to a rise-fall or fall-rise contour. This factor is referred to as *pitch contour*.

The orthogonal manipulation of these dimensions yielded 24 unique vocal melodies: 3 brightness changes (da-do/small, do-di/medium, da-di/large) x 2 pitch changes (P5, TT) x 2 brightness contours (rise-fall, fall-rise) x 2 pitch contours (rise-fall, fall-rise). In order to keep participants focused on pitch change and not absolute pitch, these 24 tone sequences were synthesized at three different pitch heights (low, middle, high). For the rise-fall pitch contour, the first tone was produced on F_3, G_3, or A_3, and for the fall-rise pitch contour, the first tone was produced on C_4, D_4, or E_4. Abstracting a step further, the combination of centroid contour and pitch contour for each tone sequence could be described as congruent or incongruent. Examples of congruent and incongruent stimuli are provided in Figure 3.

PROCEDURE

Stimuli were presented and responses recorded using custom software running on a Power Mac computer. Participants were asked to judge the interval size between the flanking notes and the central note on a 5-point scale, with "1" being a very small pitch change and "5" being a very large pitch change. Participants were encouraged to make judgments as quickly as possible (after Russo & Thompson, 2005b).

The experimental trials were blocked by pitch contour, with block order counterbalanced across participants to minimize carry-over effects. Within each pitch contour block, the trials were independently randomized in 2 sets of 36, thus yielding 2 repetitions of each stimulus and a total of 72 trials. The participants received each block twice, for a total of 4 repetitions of each stimulus (288 trials). At the beginning of the first instance of each block type, participants received either five practice trials, or as many as they needed to become familiar with the task.

Results and Discussion

Ratings of interval size were collapsed across transposition and repetition. These ratings were then analyzed using repeated measures ANOVA, with Pitch Contour (rise-fall, fall-rise), Pitch Change (P5, TT), Brightness Contour (rise-fall, fall-rise), and Brightness Change (small, medium, large) as within-subject factors. For the purpose of space and clarity, we report here confirmatory analyses meant to assess our experimental predictions, detailed above. Descriptive statistics (Appendix Table 1) and ancillary analyses are reported in the Appendix.

As expected, there was a significant main effect of Pitch Change, $F(1, 26) = 37.21$, $p < .001$, $\eta_G{}^2 = .093$, such that participants perceived P5 to be a larger interval than TT. There was also a significant main effect of Brightness Change, $F(2, 52) = 13.87$, $p < .001$, $\eta_G{}^2 = .042$. Large brightness changes received higher interval size ratings than small brightness changes, $t(26) = 4.72$, $p < .001$, $d = 0.91$, or medium brightness changes, $t(26) = 5.76$, $p < .001$, $d = 1.11$, and there was no difference in ratings between small and medium brightness changes, $t(26) = 0.36$, $p = .72$, $d = 0.07$.

As predicted, there was a significant two-way interaction between Brightness Contour and Pitch Contour, $F(1, 26) = 11.27$, $p = .002$, $\eta_G{}^2 = .020$. This interaction was driven by congruent stimuli receiving higher interval size ratings than incongruent stimuli, $t(26) =$

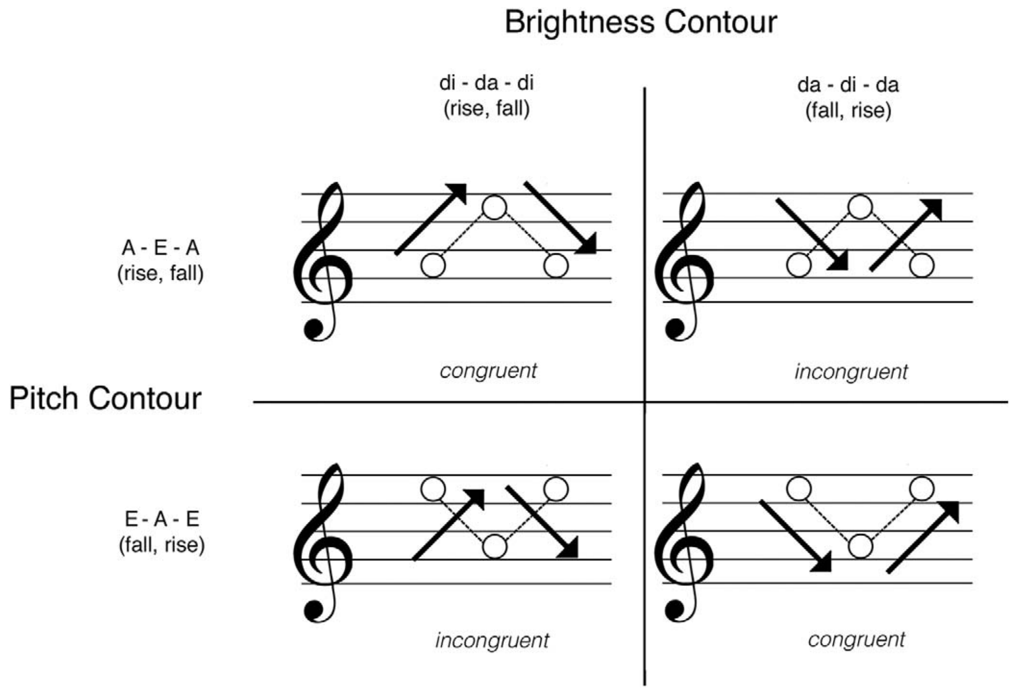

FIGURE 3. These four example stimuli demonstrate the crossing of pitch contour and brightness contour dimensions (notated in the treble clef for simplicity). The upper-left and lower-right examples are congruent. The lower-left and upper-right examples are incongruent. The pitch contour dimension has been traced using dashed lines, while the brightness contour dimension has been traced using solid arrows.

3.36, $p = .002$, $d = 0.65$. Most critically, there was a significant three-way interaction between Brightness Change, Brightness Contour, and Pitch Contour, $F(2, 52) = 5.27$, $p = .008$, $\eta_G^2 = .012$. Further analysis indicates that the effect of congruency is largest for large brightness changes, $t(26) = 4.66$, $p < .001$, $d = 0.90$, relatively smaller for medium brightness changes, $t(26) = 2.27$, $p = .03$, $d = 0.44$, and nonsignificant for small brightness changes, $t(26) = 0.01$, $p = .99$, $d = 0.003$.

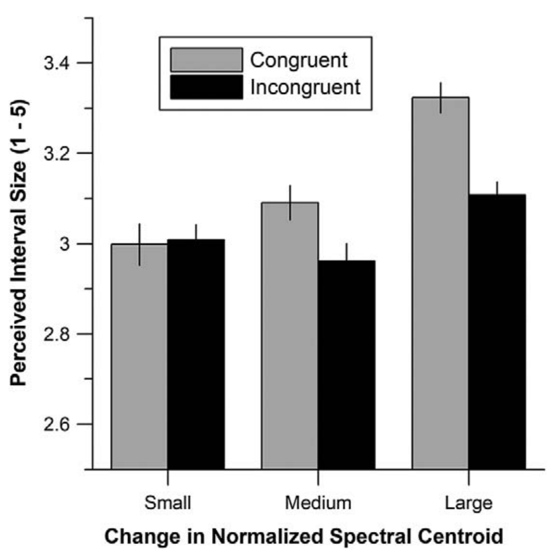

FIGURE 4. Perceived interval size (with standard error bars) for congruent and incongruent note sequences involving small, medium, and large changes in brightness.

Figure 4 plots perceived intervals size for congruent and incongruent note sequences involving small, medium, and large changes in brightness.

Experiment 1b — Control Experiment

There is a tendency across languages to produce so called "high vowels" such as [i] at a higher fundamental frequency than low vowels such as [a] (Whalen & Levitt, 1995). Although Fowler and Brown (1997, Experiment 1) found that this production effect does not hold for sung vowels, we conducted a control experiment to investigate whether the results obtained in Experiment 1a were somehow due to differences in the perceived pitch of individual vowels. Participants were required to match the frequency of a pure tone with individual syllables.

Method

PARTICIPANTS

Fourteen participants with varying levels of musical experience ($M = 2.21$ years, $SD = 3.17$) were recruited from the University of Toronto Mississauga community. These participants included 5 men and 9 women, ranging in age from 18 to 21 ($M = 18.79$ years, $SD = 0.89$). All were awarded course credit or compensated at a rate of $10 per hour for their contribution. No participants had abnormal hearing.

STIMULI

Syllables were drawn from stimuli used in Experiment 1. Specifically, we used di and da syllables at each of two pitch levels (G_3 and D_4). Pure tones were generated using Max/MSP and were randomly assigned a starting frequency that was 2, 4, or 6 semitones above or below the nominal pitch of each test syllable.

PROCEDURE

On a given trial, participants heard an alternating sequence of test syllable and pure tone. Using a mouse, participants were required to adjust the frequency of the pure tone until they perceived it to be equivalent in pitch to the syllable. Adjustments to frequency were made in increments or decrements of 1 Hz.

Results and Discussion

Normalized error was calculated for each trial by subtracting the pitch-matched frequency of the pure tone from the frequency of the syllable and dividing by the frequency of the target. These data were submitted to a one-way ANOVA. No effect of syllable was found, $F(1, 26) = 1.55$, $p = .22$, $\eta_G^2 = .056$. The results of the control experiment support the view that brightness effects observed in Experiment 1 were due to alterations in the perception of relative pitch rather than systematic distortions of absolute pitch.

Experiment 2 — Contour Judgments

If syllable content is able to influence perception of interval size, it may also extend to other aspects of relative pitch processing. In Experiment 2, we investigated whether changes in brightness arising from vowel content can influence the perception of contour. In each trial, participants heard a two-note sequence with rising or falling pitch contour that was synthesized using the same vocal synthesis parameters used in Experiment 1. Our predictions and analysis plan for Experiment 2 were pre-registered and can be viewed at https://osf.io/9xf43.

Method

PARTICIPANTS

Twenty participants with varying levels of music experience ($M = 2.55$ years, $SD = 4.49$) were recruited from the Ryerson University community. These participants included 2 men and 18 women, ranging in age from 18 to 24 ($M = 20.69$ years, $SD = 2.06$). All were awarded course credit or compensated at a rate of $10 per hour for their contribution. No participants reported abnormal hearing.

STIMULI

Two-note sequences were created by altering the stimuli used in Experiment 1. The final syllable was dropped from each sequence (e.g., di-da-di became di-da).

PROCEDURE

On each trial, participants heard a two-note sequence and were asked to categorize the pitch direction as rising or falling. Participants were encouraged to make their responses as quickly as possible.

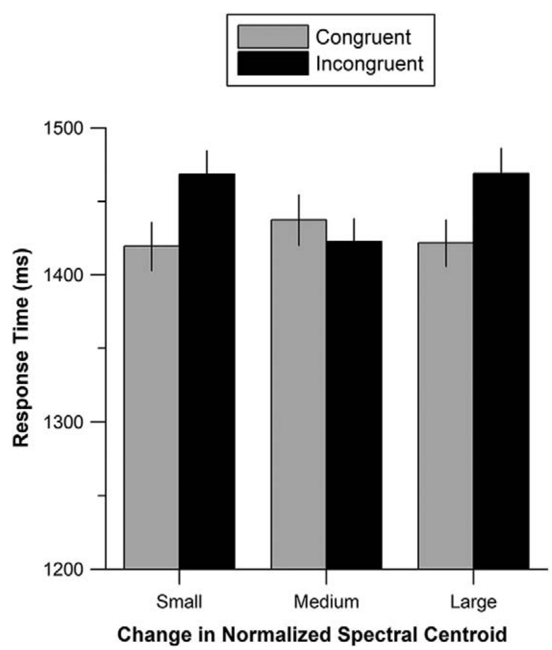

FIGURE 5. Response times (with standard error bars) for congruent and incongruent note sequences involving small, medium, and large changes in brightness.

Results and Discussion

The accuracy of contour judgments was high ($M = 0.83$, $SD = 0.18$) and its distribution was negatively skewed (-.184), indicative of a ceiling effect (Bulmer, 1979). Accordingly, our analyses focused on response time data only. Figure 5 plots response times for congruent and incongruent trials when brightness changes are small, medium, and large.

Response times were discarded if they were more than 3 standard deviations from the mean response time across participants. Next, response times were collapsed across pitch change, transposition, and repetition. These response times were then analyzed using repeated measures ANOVA, with Pitch Contour (rise-fall, fall-rise), Brightness Contour (rise-fall, fall-rise), and Brightness Change (small, medium, large) as within-subjects factors.

As predicted, there was a significant interaction between Pitch Contour and Brightness Contour, $F(1, 19) = 4.50$, $p = .05$ $\eta_G^2 = .002$. This was driven by faster responses to congruent than incongruent trials, $t(19) = 2.12$, $p = .05$, $d = .47$. Most critically, there was a significant three-way interaction between Brightness Change, Brightness Contour, and Pitch Contour, $F(2, 38) = 3.87$, $p = .03$, $\eta_G^2 = .003$. Further analysis focused on each brightness change condition separately. For large brightness changes, the effect of congruency was significant, $t(19) = 2.24$, $p = .04$, $d = 0.50$, with faster responses for congruent than incongruent responses. For medium brightness changes, the effect of congruency was nonsignificant, $t(19) = 1.15$, $p = .26$, $d = 0.26$. Finally for small brightness changes, the effect of congruency was significant, $t(19) = 2.14$, $p = .05$, $d = 0.48$, with faster responses for congruent than incongruent responses.

General Discussion

The results of Experiments 1a demonstrated that note-to-note changes in vowel content can influence perception of interval size in vocal melodies. This striking finding may be interpreted as a novel extension of our prior work on the influence of brightness on perceived interval size, obtained using synthetic and instrument timbres (Russo & Thompson, 2005b). Experiment 1b confirmed that these results were not due to systematic pitch distortions. Experiment 2 showed that note-to-note changes in brightness may also influence judgments of pitch contour.

In articulatory phonetics, intrinsic pitch relates to the pitch height at which the vowel tends to be produced. With regard to the vowels considered here, a "high" vowel like [i] has a higher intrinsic pitch than a "low" vowel like [a], with [o̜] falling somewhere in-between. The higher intrinsic pitch of [i] may have a basis in pharyngeal wall expansion (Ewan, 1975; Titze, 2008) and/or a constriction of the laryngeal muscles (Dhyr, 1990). From a perceptual stand point, there is some evidence that a high vowel will be heard as lower in pitch than an open vowel that is produced at the same fundamental frequency (Fowler & Jones, 1997). Although there was no such finding here in the context of pitch matching (1B), we do find convergent results in our interval size task (1A).

So, how exactly does this prior work on the intrinsic pitch of vowels reconcile with the current findings that have been characterized with respect to the brightness of vowels? The brightness of vowels was defined here on the basis of the normalized spectral centroid but it may have been similarly derived as the inverse of F1, or some other dimension that has been used to characterize the vowel articulatory space (e.g., F2–F1). We found that tone sequences that were congruent in brightness change and pitch change led to larger judgments of interval size and quicker judgments of interval contour than those that were incongruent. The extent to which these effects generalize across the articulatory vowel space and for vowels as they are naturally produced in real singing, remains to be determined. Having a more fulsome sampling of vowels drawn from across this space may allow future work to untangle any unique variance that may be attributable to different vocal-acoustic parameters.

The findings of this study may also have implications for text setting. The primary consideration for text setting tends to be lyric intelligibility (e.g., Fine & Ginsborg, 2014; Johnston, Huron, & Collister, 2014). Syllabic setting (one note per syllable) is said to be more intelligible than melismatic setting (more than one note per syllable). A secondary consideration has been the influence of intrinsic stress properties of syllables on rhythm. On the basis of the current investigation, it appears that the influence of vowel content on melodic pitch perception is yet another factor to consider in text setting.

While our focus here has been on the perception of pitch relations in vocal melodies, future work might also consider note-to-note expectations. Music theorists have often asserted that the size of a melodic interval has special significance for expectancy about the next note to follow (e.g., Huron, 2006; Larson, 2002; Margulis, 2005; Meyer, 1956; Narmour, 1990). If vowel content exerts a subtle but systemic influence on pitch relations between notes as demonstrated here, it should also influence melodic expectancies, yielding rational consequences for tension and arousal.

Author Note

We thank Fran Copelli for assistance with data collection and preparation of figures. Funding for this study was provided by Natural Science and Engineering Research Council Discovery Grants awarded to the first and third authors.

Correspondence concerning this article should be addressed to Frank A. Russo, Department of Psychology, Ryerson University, Toronto, Canada, M5B 2K3. E-mail: russo@ryerson.ca

References

ABEL, M. K., LI, H. C., RUSSO, F. A., SCHLAUG, G., & LOUI, P. (2016). Audiovisual interval size estimation is associated with early musical training. *PLoS ONE*, *11*, e0163589. https://doi.org/10.1371/journal.pone.0163589

ALLEN, E. J., & OXENHAM, A. J. (2014). Symmetric interactions and interference between pitch and timbre. *Journal of the Acoustical Society of America*, *135*, 1371–1379. https://doi.org/10.1121/1.4863269

BULMER, M. G. 1979. *Principles of statistics*. New York: Dover.

DHYR, N. (1990). The activity of the cricothyroid muscle and the intrinsic fundamental frequency in Danish vowels. *Phonetica, 47*, 141–154.

EWAN, W. G. (1975). Explaining the intrinsic pitch of vowels. *Journal of the Acoustical Society of America*, *58*, S40–S40.

FINE, P. A., & GINSBORG, J. (2014). Making myself understood: Perceived factors affecting the intelligibility of sung text. *Frontiers in Psychology*, *5*, 809. https://doi.org/10.3389/fpsyg.2014.00809

FOWLER, C. A. & BROWN, J.M. (1997). Intrinsic f0 differences in spoken and sung vowels and their perception by listeners. *Perception and Psychophysics, 59,* 729–738. https://doi.org/10.3758/BF03206019

HALL, M. D., & BEAUCHAMP, J. W. (2009). Clarifying spectral and temporal dimensions of musical instrument timbre. *Canadian Acoustics*, *37*(1), 3–22.

HILLENBRAND, J., GETTY, L. A., CLARK, M. J., & WHEELER, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, *97*, 3099–3111.

HURON, D. B. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: MIT Press.

JOHNSON, R., HURON, D., & COLLISTER, L. B. (2014). Music and lyrics interactions and their influence on recognition of sung words: an investigation of word frequency, rhyme, metric stress, vocal timbre, melisma, and repetition priming. *Empirical Musicology Review. 9*, 2–20.

KAE LABS. (2005). *VocalWriter* [Software]. Woodinville, WA: KAE Labs.

KLATT, D. H., & KLATT, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America*, *87*, 820–857. https://doi.org/10.1121/1.398894

LARSON, S. (2002). Musical forces, melodic expectation, and jazz melody. *Music Perception*, *19*, 351–385.

LIVINGSTONE, S. R., CHOI, D., & RUSSO, F. A. (2014). The influence of vocal training and acting experience on measures of voice quality and emotional genuineness. *Frontiers in Psychology*, *5*, 156. https://doi.org/10.3389/fpsyg.2014.00156

MARGULIS, E. H. (2005). A model of melodic expectation. *Music Perception*, *22*, 663–714.

MCADAMS, S., & GIORDANO, B. L. (2011). The perception of musical timbre. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (pp. 72–80). Oxford, UK: Oxford University Press.

MCDERMOTT, J. H., LEHR, A. J., & OXENHAM, A. J. (2008). Is relative pitch specific to pitch? *Psychological Science*, *19*, 1263–1271. https://doi.org/10.1111/j.1467-9280.2008.02235.x

MEYER, L. B. (1956). *Emotion and meaning in music*. Chicago, IL: University of Chicago Press.

NARMOUR, E. (1990). *The analysis and cognition of basic melodic structures: The implication-realization model*. Chicago, IL: University of Chicago Press.

RUSSO, F. A., & THOMPSON, W. F. (2005a). The subjective size of melodic intervals over a two-octave range. *Psychonomic Bulletin and Review*, *12*, 1068–1075. https://doi.org/10.3758/BF03206445

RUSSO, F. A., & THOMPSON, W. F. (2005b). An interval size illusion: The influence of timbre on the perceived size of melodic intervals. *Perception and Psychophysics*, *67*, 559–568. https://doi.org/10.3758/BF03193514

SCHUBERT, E., & WOLFE, J. (2006). Does timbral brightness scale with frequency and spectral centroid? *Acta Acustica United with Acustica*, *92*(5), 820–825.

SLAWSON, W. (1985). *Sound color*. London, UK: University of California Press.

SUNDBERG, J. (1972). A perceptual function of the singing formant. *Speech Transmission Laboratory Quarterly Progress and Status Report*, *2*, 61–63.

SUNDBERG, J. (1994). Perceptual aspects of singing. *Journal of Voice*, *8*, 106–122. https://doi.org/10.1016/S0892-1997(05)80303-0

Titze, I. R. (2008). Nonlinear source–filter coupling in phonation: Theory. *Journal of the Acoustical Society of America*, *123*, 1902–1915.

Thompson, W. F., Russo, F. A., & Livingstone, S. R. (2010). Facial expressions of singers influence perceived pitch relations. *Psychonomic Bulletin and Review*, *17*, 317-322.

Thompson, W. F., Graham, P., & Russo, F. A. (2005). Seeing music performance: Visual influences on perception and experience. *Semiotica*, *156*, 203–227. https://doi.org/10.1515/semi.2005.2005.156.203

Thompson, W., Peter, V., Olsen, K. N., & Stevens, C. J. (2012). The effect of intensity on relative pitch. *Quarterly Journal of Experimental Psychology*, *65*(10), 2054–2072. https://doi.org/10.1080/17470218.2012.678369

Whalen, D. H., & Levitt, A. G. (1995). The universality of intrinsic F0 of vowels. *Journal of Phonetics*, *23*(3), 349–366. https://doi.org/10.1016/S0095-4470(95)80165-0

Appendix

TABLE 1. *Experiment 1 Mean Ratings*

Main effect: Pitch Change

Pitch Change	Mean [*SD*]
P5	3.27 [1.06]
TT	2.99 [1.06]

Main effect: Brightness Change

Brightness Change	Mean [*SD*]
Small	3.05 [1.10]
Medium	3.07 [1.03]
Large	3.26 [1.07]

Interaction: Pitch Contour x Brightness Contour

Pitch Contour	Brightness Contour	Mean [*SD*]
Fall-Rise	Fall-Rise	3.19 [1.08]
Fall-Rise	Rise-Fall	3.03 [1.04]
Rise-Fall	Fall-Rise	3.09 [1.05]
Rise-Fall	Rise-Fall	3.18 [1.10]

Interaction: Brightness Change x Pitch Contour x Brightness Contour

Brightness Change	Pitch Contour	Brightness Contour	Mean [*SD*]
Small	Fall-Rise	Fall-Rise	3.11 [1.15]
Small	Fall-Rise	Rise-Fall	3.01 [1.01]
Small	Rise-Fall	Fall-Rise	3.10 [1.08]
Small	Rise-Fall	Rise-Fall	2.99 [1.14]
Medium	Fall-Rise	Fall-Rise	3.11 [0.98]
Medium	Fall-Rise	Rise-Fall	3.02 [1.06]
Medium	Rise-Fall	Fall-Rise	2.98 [1.02]
Medium	Rise-Fall	Rise-Fall	3.17 [1.03]
Large	Fall-Rise	Fall-Rise	3.36 [1.07]
Large	Fall-Rise	Rise-Fall	3.07 [1.04]
Large	Rise-Fall	Fall-Rise	3.21 [1.04]
Large	Rise-Fall	Rise-Fall	3.39 [1.09]

Experiment 1a

RESULTS: RATINGS

Ratings of interval size were collapsed across transposition and repetition. These ratings were then analyzed using repeated measures ANOVA, with Pitch Contour (rise-fall, fall-rise), Pitch Change (P5, TT), Brightness Contour (rise-fall, fall-rise), and Brightness Change (small, medium, large) as within-subjects factors.

There was a significant main effect of Pitch Change, $F(1, 26) = 13.87$, $p < .001$, $\eta_G^2 = .093$. There was also a significant main effect of Brightness Change, $F(2, 52) = 37.21$, $p < .001$, $\eta_G^2 = .042$. All other main effects were nonsignificant.

There was a significant two-way interaction between Pitch Contour and Pitch Change, $F(1, 26) = 28.59$, $p < .001$, $\eta_G^2 = .024$. There was a significant two-way interaction between Pitch Contour and Brightness Contour, $F(1, 26) = 11.27$, $p = .002$, $\eta_G^2 = .020$. There was a significant two-way interaction between Brightness Contour and Brightness Change, $F(2, 52) = 4.01$, $p = .024$, $\eta_G^2 = .005$. All other two-way interactions were nonsignificant.

There was a significant three-way interaction between Pitch Contour, Pitch Change, and Brightness Change, $F(2, 52) = 3.97$, $p = .025$, $\eta_G^2 = .005$. There was a significant three-way interaction between Brightness Change, Brightness Contour, and Pitch Contour, $F(2, 52) = 5.27$, $p = .008$, $\eta_G^2 = .012$. All other three-way interactions were nonsignificant.

RESULTS: RESPONSE TIMES

Response times were collapsed across transposition and repetition. These response times were then analyzed using repeated measures ANOVA, with Pitch Contour (rise-fall, fall-rise), Pitch Change (P5, TT), Brightness Contour (rise-fall, fall-rise), and Brightness Change (small, medium, large) as within-subjects factors.

There was a main effect of Pitch Contour, $F(1, 26) = 4.50$, $p = .04$, $\eta_G^2 = .001$. All other main effects were nonsignificant.

There was also a significant three-way interaction between Pitch Contour, Brightness Contour, and Brightness Change, $F(2, 52) = 4.19$, $p = .021$, $\eta_G^2 = .003$. All other interactions were nonsignificant.

Experiment 2

RESULTS: RESPONSE TIMES

Response times were discarded if they were more than 3 standard deviations from the mean response time across participants. Next, response times were collapsed across pitch change, transposition, and repetition. These response times were then analyzed using repeated measures ANOVA, with Pitch Contour (rise-fall, fall-rise), Brightness Contour (rise-fall, fall-rise), and Brightness Change (small, medium, large) as within-subjects factors.

There was a main effect of Brightness Contour, $F(1, 19) = 5.30$, $p = .03$, $\eta_G^2 = .002$. All other main effects were nonsignificant.

There was a significant two-way interaction between Pitch Contour and Brightness Contour, $F(1, 19) = 4.50$, $p = .05$, $\eta_G^2 = .002$. There was also a significant two-way interaction between Pitch Contour and Brightness Change, $F(2, 38) = 4.69$, $p = .02$, $\eta_G^2 = .004$. Finally, there was a significant two-way interaction between Brightness Contour and Brightness Change, $F(2, 38) = 16.57$, $p < .001$, $\eta_G^2 = .012$.

Lastly, there was a significant three-way interaction between Pitch Contour, Brightness Contour, and Brightness Change, $F(2, 38) = 3.87$, $p = .03$, $\eta_G^2 = .003$. All other interactions were nonsignificant.