ANALYSIS OF INTONATION TRAJECTORIES IN SOLO SINGING

Jiajie Dai, Matthias Mauch, Simon Dixon
Centre for Digital Music, Queen Mary University of London, United Kingdom
{j.dai, m.mauch, s.e.dixon}@qmul.ac.uk

ABSTRACT

We present a new dataset for singing analysis and modelling, and an exploratory analysis of pitch accuracy and pitch trajectories. Shortened versions of three pieces from *The Sound of Music* were selected: "Edelweiss", "Do-Re-Mi" and "My Favourite Things". 39 participants sang three repetitions of each excerpt without accompaniment, resulting in a dataset of 21762 notes in 117 recordings. To obtain pitch estimates we used the *Tony* software's automatic transcription and manual correction tools. Pitch accuracy was measured in terms of pitch error and interval error. We show that singers' pitch accuracy correlates significantly with self-reported singing skill and musical training. Larger intervals led to larger errors, and the tritone interval in particular led to average errors of one third of a semitone. Note duration (or inter-onset-interval) had a significant effect on pitch accuracy, with greater accuracy on longer notes. To model drift in the tonal centre over time, we present a sliding window model which reveals patterns in the pitch errors of some singers. Based on the trajectory, we propose a measure for the magnitude of drift: tonal reference deviation (TRD). The data and software are freely available. [1]

1. INTRODUCTION

Singing is common in all human societies [2], yet the factors that determine singing proficiency are still poorly understood. Many aspects are important to singing, including pitch, rhythm, timbre, dynamics and lyrics; here we focus entirely on the pitch dimension. Music psychologists have studied singing pitch [4, 6, 18], and engineers have developed advanced software for automatic pitch tracking [5, 11, 21], but the process of annotating and analysing the pitch of singing data remains a laborious task. In this paper, we present a new extensive dataset for the analysis of unaccompanied solo singing, complete with audio, pitch tracks, and hand-annotated note tracks matched to the scores of the music. In addition, we provide an analysis of the data with a focus on intonation: pitch errors, interval errors, pitch drift, and the factors that influence these phenomena.

Intonation, defined as "accuracy of pitch in playing or singing" [23], or "the act of singing or playing in tune" [12], is one of the main priorities in choir rehearsals [9] and in choral practice manuals (e.g. [3]). Good intonation involves the adjustment of pitch to maximise the consonance of simultaneous notes, but it also has a temporal aspect, particularly in the absence of instrumental accompaniment, where the initial tonal reference can be forgotten over time [15]. A cappella ensembles frequently observe a change in tuning over the duration of a piece, even when they are unable to detect any local changes. This phenomenon, called *intonation drift* or pitch drift [22], usually exhibits as a lowering of pitch, or downward drift [1]. Several studies present evidence that drift is induced by harmonic progressions as singers negotiate the tradeoff between staying in tune and singing in just intonation [7,10,24]. Yet this is not the only cause of drift, since drift is also observed in solo singing, such as unaccompanied solo folk songs [17] and even queries to query-by-humming systems [20]. A factor that has received relatively little attention in the singing research community is the effect of note duration on singing accuracy [8], so one of our aims in this paper is to explore the effect of duration.

The definitions of intonation given above imply the existence of a reference pitch, which could be provided by accompanying instruments or (as in the present case) could exist solely in the singer's memory. This latter case allows for the reference to change over time, and thus explain the phenomenon of drift. We introduce a novel method to model this internal reference as the pitch which minimises the intonation error given some weighted local context, and we compare various context windows for parametrising our model. Using this model of reference pitch, we compute pitch error as the signed pitch difference relative to the reference pitch and score, measured in semitones on an equal-tempered scale. Interval error is measured on the same scale, without need of any reference pitch, and pitch drift is given by the trajectory of score-normalised reference pitch over time.

In this paper we explore which factors may explain intonation error in our singing data. The effects of four singer factors, obtained by self-report, were tested for significance. Most of the participants in this study were amateur singers without professional training. Their musical background, years of training, frequency of practice and self-reported skill were all found to have a significant effect on

[1] see Data Availability, Section 7

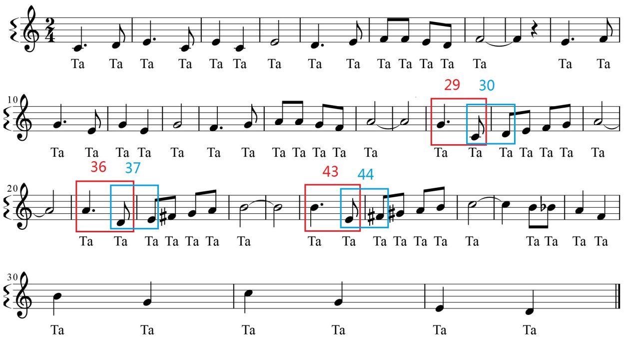


Figure 1: Score of piece Do-Re-Mi, with some intervals marked (see Section 3)

Table 1: Summary details of the three songs used in this study.

Title	Tempo (BPM)	Key	Notes
Edelweiss	80	B♭	54
Do-Re-Mi	120	C	59
My Favourite Things	132	Em	73

intonation errors. We then considered as piece factors three melodic features, note duration, interval size and the presence of a tritone interval, for their effect on intonation. All of these features had a significant effect on both pitch and interval error. Finally we consider the pitch drift trajectories of individual singers. Our model tracks the direction and magnitude of cumulative pitch errors and captures how well participants remain in the same key. Some trajectories have periodic structure, revealing systematic errors in the singing.

2. MATERIALS AND METHODS

2.1 Musical material

We chose three songs from the musical "The Sound of Music" as our material: "Edelweiss", "Do-Re-Mi" (shown in Figure 1) and "My Favourite Things." Despite originating from one work, the pieces were selected as being diverse in terms of tonal material and tempo (Table 1), well-known to many singers, and yet sufficiently challenging for amateur singers. The pieces were shortened so as to contain a single verse without repeats, which the participants were asked to sing to the syllable "ta". In order to observe long-term pitch trends, each song was sung three times consecutively. Each trial lasted a little more than 5 minutes.

2.2 Participants

We recruited 39 participants (12 male, 27 female), most of whom are members of our university's music society or our music-technology focused research group. Some participants took part in the experiments remotely. The age of the participants ranged from 20 to 27 years (mean 23.3, median 23 years). We asked all participants to self-assess their musical background with questions loosely based on the Goldsmiths Musical Sophistication Index [16].[2] Table 2 shows the results, suggesting a range of skill levels, with a strong bias towards amateur singers.

Table 2: Self-reported musical experience

Musical Background		Instrumental Training	
None	5	None	5
Amateur	27	1–2 years	15
Semi-professional	5	3–4 years	7
Professional	2	5+ years	12
Singing Skill		**Singing Practice**	
Poor	2	None	4
Low	25	Occasionally	22
Medium	9	Often	12
High	3	Frequently	1

2.3 Recording procedure

Participants were asked to sing each piece three times on the syllable 'ta'. They were given the starting note but no subsequent accompaniment, except unpitched metronome clicks.

2.4 Annotation

We used the software *Tony*[3] to annotate the notes in the audio files [13]: pitch track and notes were extracted using the pYIN algorithm [14] and then manually checked and, if necessary, corrected. Approximately 28 corrections per recording were necessary; detailed correction metrics on this data have been reported elsewhere [13].

2.5 Pitch metrics

The *Tony* software outputs the median fundamental frequency f_0 for every note. We relate fundamental frequency to musical pitch p as follows:

$$p = 69 + 12 \log_2 \frac{f_0}{440\,\text{Hz}} \tag{1}$$

This scale is chosen such that a difference of 1 corresponds to 1 semitone. For integer values of p the scale coincides with MIDI pitch numbers, with reference pitch A4 tuned to 440 Hz ($p = 69$).

2.5.1 Interval Error

A musical interval is the difference between two pitches [19] (which is proportional to the logarithm of the ratio of the fundamental frequencies of the two pitches). Using Equation 1, we define the interval from a pitch p_1 to the pitch p_2 as $i = p_2 - p_1$ and hence we can define the interval error between a sung interval i and the expected nominal interval i_n (given by the musical score) as:

$$e^{int} = i - i_n \tag{2}$$

[2] The questions were: How do you describe your musical background? How many years do you have instrument training? How do you describe your singing skills? How often do you practice your singing skills?

[3] https://code.soundsoftware.ac.uk/projects/tony

Hence, for a piece of music with M intervals $\{e_1^{int}, \ldots, e_M^{int}\}$, the mean absolute interval error (MAIE) is calculated as follows:

$$\text{MAIE} = \frac{1}{M}\sum_{i=1}^{M} |e_i{}^{int}| \tag{3}$$

2.5.2 *Tonal reference curves and pitch error*

In unaccompanied singing, pitch error is ill-defined, since singers use intonation with respect to their internal reference, which we cannot track directly. If it is assumed that this internal reference doesn't change, we can estimate it via the mean error with respect to a nominal (or given) reference pitch. However, it is well-known that unaccompanied singers (and choirs) do not maintain a fixed internal reference (see Section 1). Previously, this has been addressed by estimating the singer's reference frequency using linear regression [15], but as there is no good reason to assume that drift is linear, we adopt a sliding window approach in order to provide a local estimate of tuning reference.

The first step is to take the annotated musical pitches p_i of a recording and remove the nominal pitch s_i given by the score, $t_i^* = p_i - s_i$, which we adjust further by subtracting the mean: $t_i = t_i^* - \bar{t}^*$. The resulting raw tonal reference estimates t_i are then used as a basis for our tonal reference curves and pitch error calculations.

The second step is to find a smooth trajectory based on these raw tonal reference estimates. For each note, we calculate the weighted mean of t_i in a context window around the note, obtaining the reference pitch c_i, from which the pitch error can be calculated:

$$c_i = \sum_{k=-n}^{n} w_k t_{i+k}, \tag{4}$$

where $\sum_{k=-n}^{n} w_k = 1$. Any window function $W = \{w_k\}$ can be used in Equation 4. We experimented with symmetric windows with two different window shapes (rectangular and triangular) and seven window sizes (3, 5, 7, 9, 11, 15 and 25 notes) to arrive at smooth tonal reference curves. The rectangular window $W^{R,N} = \{w_k^{R,N}\}$ centred at the i^{th} note is used to calculate the mean of its N-note neighbourhood, giving the same weight to all notes in the neighbourhood, but excluding the i^{th} note itself:

$$w_k^{R,N} = \begin{cases} \frac{1}{N-1}, & 1 \leq |k| \leq \frac{N-1}{2} \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

The triangular window $W^{T,N} = \{w_k^{T,N}\}$ gives more weight to notes near the i^{th} note (while still excluding the i^{th} note itself). For example, if the window size is 5, then the weights are proportional to 1, 2, 0, 2, 1. More generally:

$$w_k^{T,N} = \begin{cases} \frac{2N+2-4|k|}{N^2-1}, & 1 \leq |k| \leq \frac{N-1}{2} \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

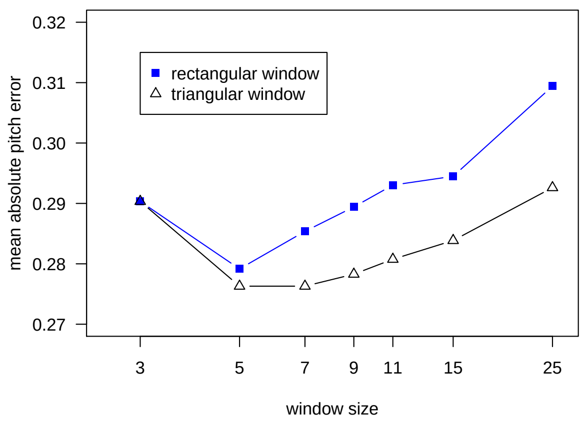


Figure 2: Pitch error (MAPE) for different sliding windows.

The smoothed tonal reference curve c_i is the basis for calculating the pitch error:

$$e_i^p = t_i - c_i, \tag{7}$$

so for a piece with M notes with associated pitch errors $e_1^p, \ldots, e_M^p$, the mean absolute pitch error (MAPE) is:

$$\text{MAPE} = \frac{1}{M}\sum_{i=1}^{M} |e_i^p|. \tag{8}$$

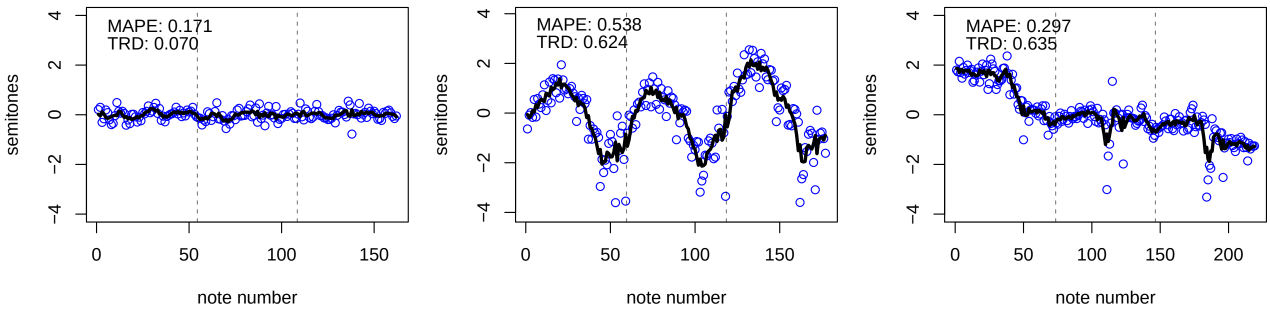
2.5.3 *Tonal reference deviation*

The tonal reference curves c_i can also be used to calculate a new measure of the extent of fluctuation of a singer's reference pitch. We call this measure tonal reference deviation (TRD), calculated as the standard deviation:

$$\text{TRD} = \sqrt{\frac{1}{M-1}\sum_{i=1}^{M}(c_i - \bar{c}_M)^2}. \tag{9}$$

3. RESULTS

We first compare multiple choices of window for the calculation of the smoothed tonal reference curves c_i (Section 2.5.2), which provide the local tonal reference estimate used for calculating mean absolute pitch error (MAPE). We assume that the window that gives rise to the lowest MAPE models the data best. Figure 2 shows that for both window shapes an intermediate window size N of 5 notes minimises MAPE, with the triangular window working best (MAPE $= 0.276$ semitones, computed over all singers and pieces). Hence, we use this window for all further investigations relating to pitch error, including tonal reference curves, and for understanding how pitch error is linked to note duration and singers' self-reported skill and experience.



(a) Edelweiss, singer 11 (b) Do-Re-Mi, singer 39 (c) My Favourite Things, singer 31

Figure 3: Examples of tonal reference trajectories. Dashed vertical lines delineate the three repetitions of the piece.

3.1 Smoothed tonal reference curves

The smoothed curves exhibit some unexpected behaviour. Figure 3 shows three examples of different participants and pieces. Several patterns emerge. Figure 3a shows a performance in which pitch error is kept within half a semitone and tonal reference is almost completely stable. This is reflected in very low values of MAPE (0.171) and TRD (0.070), respectively. However, most singers' tonal reference curves fluctuate. For example, Figure 3b illustrates a tendency of some singers to smoothly vary their pitch reference in direct response to the piece. The trajectory shows a periodic structure synchronised with the three repetitions of the piece. The fluctuation measure TRD is much higher as a result (0.624). This is a common pattern we have observed. The third example (Figure 3c) illustrates that strong fluctuations are not necessarily periodic. Here, TRD (0.635) is nearly identical, but originates from a mostly consistent downward trajectory. The singer makes significant errors in the middle of each run of the piece, most likely due to the difficult interval of a downward tritone occurring twice (notes 42 and 50; more discussion below). Comparing Figures 3b and 3c also shows that MAPE and TRD are not necessarily related. Despite large fluctuations (TRD) in both, pitch error (MAPE) is much smaller in Figure 3c (0.297).

Turning from the trajectories to pitch error measurements, we observe that the three pieces show distinct patterns (Figure 4). The first piece, Edelweiss, appears to be the easiest to sing, with relatively low median pitch errors. In Do-Re-Mi, the third quarter of the piece appears much more difficult than the rest. This is most likely due to faster runs and the presence of accidentals, taking the singer out of the home tonality. Finally, My Favourite Things exhibits a very distinct pattern, with relatively low pitch errors throughout, except for one particular note (number 50), which is reached via a downward tritone, a difficult interval to sing. The same tritone (A-D♯) occurs at note 42, where the error is smaller and notably in the opposite direction (this D♯ is flat, while note 50 is over a semitone sharp on average). It appears that singers are drawn towards the more consonant (and more common) perfect fifth and fourth intervals, respectively.

	Estimate	Std. Err.	t	p
(intercept)	0.374	0.012	32.123	0.000
nominal duration	-0.073	0.004	-17.487	0.000
prev. nom. IOI	-0.021	0.004	-4.646	0.000
abs(nom. interval)	0.016	0.001	13.213	0.000
abs(next nom. interval)	0.010	0.001	8.471	0.000
tritone	0.370	0.019	19.056	0.000
quest. score	-0.011	0.001	-9.941	0.000

(a) MAPE

	Estimate	Std. Err.	t	p
(intercept)	0.481	0.015	33.124	0.000
nominal duration	-0.076	0.005	-14.570	0.000
prev. nom. IOI	-0.050	0.006	-8.984	0.000
abs(nom. interv.)	0.030	0.002	19.700	0.000
abs(next nom. interv.)	-0.006	0.002	-3.826	0.000
tritone	0.373	0.024	15.404	0.000
quest. score	-0.012	0.001	-8.665	0.000

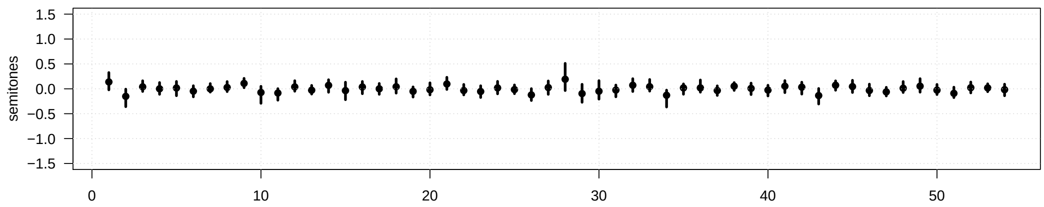
(b) MAIE

Table 3: Effects of multiple covariates on error for a linear model. t denotes the test statistic. The p value rounds to zero in all cases, indicating statistical significance.

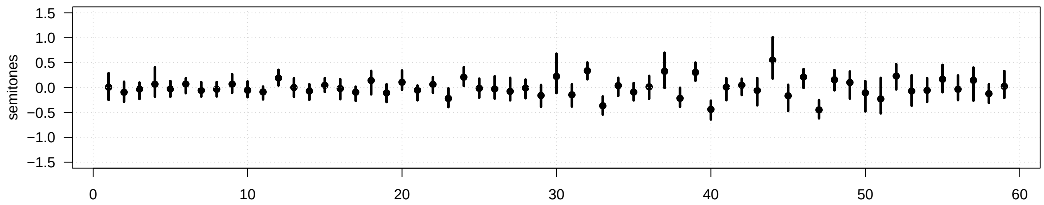
3.2 Duration, interval and proficiency factors

The observations on pitch error patterns suggest that note duration and the tritone interval may have significant impact on pitch error. In order to investigate their impact we make use of a linear model, taking into account furthermore the size of the intervals sung and singer bias via considering the singers' self assessment.

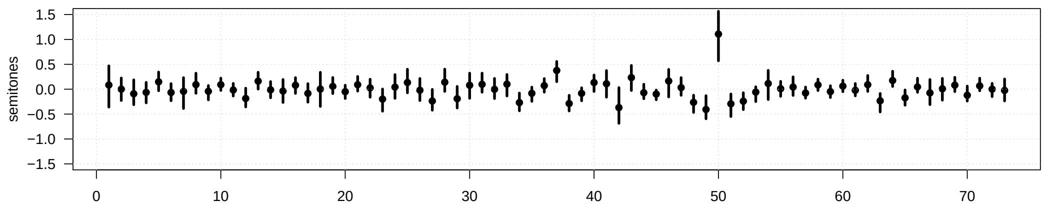
Table 3a lists all dependent variables, estimates of their effects and indicators of significance. In the following we will simply speak of how these variables influence, reduce or add to error, noting that our model gives no indication of true causation, only of correlation. We turn first to the question of whether note duration influences pitch error. The intuition is that longer notes, and notes with a longer preparation time (previous inter onset interval, IOI), should be sung more correctly. This is indeed the case. We observe a reduction of pitch error of 0.073 semitones per added second of duration. The IOI between previous and current note also reduces pitch error, but by a smaller factor (0.021 semitones per second). Conversely, absolute nominal interval size adds to absolute pitch error, by about 0.016 semitones per interval-semitone, as does



(a) Edelweiss



(b) Do-Re-Mi



(c) My Favourite Things

Figure 4: Pitch errors by note for each of the three pieces. The plots show the median values with bars extending to the first and third quartiles.

the absolute size of the next interval (0.010 semitones). The intuition about the tritone interval is confirmed here, as the presence of any tritone (whether upward or downward) adds 0.370 semitones—on average—to the absolute pitch error. The last covariate, questionnaire score, is the sum of the points obtained from the four self-assessment questions, with values ranging between 5 and 14. The result shows that there is correlation between the singers' self-assessment and their absolute pitch error. For every additional point in the score their absolute pitch error is reduced by 0.012 semitones. The picture is very similar as we do the same analysis for absolute interval error (Table 3b): the effect directions of the variables are the same.

4. DISCUSSION

We have investigated how note length relates to singing accuracy, finding that notes are sung more accurately as the singer has more time to prepare and sing them. Yet it is not entirely clear what this improvement is based upon. Do longer notes genuinely give singers more time to find the pitch, or is part of the effect we observe due to measurement or statistical artefacts? To find out, we will need to examine pitch at the sub-note level, taking vibrato and note transitions into account. Conversely, studying the effect of melodic context on the underlying pitch track could shed light on the physical process of singing, and could be used for improved physical modelling of singing.

Overall, the absolute pitch error of singers (mean: 28 cents; median: 18; std.dev.: 36) and the absolute interval error (mean: 34 cents; median: 22; std.dev.: 46) are slightly higher than those reported elsewhere [15], but this may reflect the greater difficulty of our musical material in comparison to "Happy Birthday". We also did not exclude singers for their pitch errors, although the least accurate singers had MAPE and MAIE values of more than half a semitone, i.e. they were on average closer to an erroneous note than to the correct one. That the values of MAIE and MAPE are similar is to be expected, as interval error is the limiting case of pitch error, using a minimal window containing only the current and previous note.

We used a symmetric window in this work, but this could easily be replaced with a causal (one-sided) window [15], which would also be more plausible psychologically, as the singer's internal pitch reference in our model is based equally on past sung notes and future not-yet-sung notes. However, for post hoc analysis, the fuller context might reveal more about the singer's internal state (which must influence the future tones) than the more restricted causal model.

Figure 4 shows how the three pieces in our data differ in terms of pitch accuracy. It is interesting to see that accidentals (which result in a departure from the established key), and the tritone as a particular example, seem to have a strong adverse impact on accuracy. To compile more detailed statistical analyses like the ones in Table 3 one could conduct singing experiments on a wider range of intervals, isolated from the musical context of a song. In future work we also intend to explore the interaction between singers as they negotiate a common tonal reference.

Finally, we would like to mention that some singers took prolonged breaks between runs in a three-run rendition of a song. The recording was stopped, but no new reference note was played, so the singers resumed with the memory of what they last sung. As part of the reproducible code package (see Section 7) we provide information on which recordings were interrupted and at which break. We found that the regression coefficients (Tables 3b and 3a) did not substantially change as a result of these interruptions.

5. CONCLUSIONS

We have presented a new dataset for singing analysis, investigating the effects of singer and piece factors on the intonation of unaccompanied solo singers. Pitch accuracy was measured in terms of pitch error and interval error. We introduced a new model of tonal reference computed using the local neighbourhood of a note, and found that a window of two notes each side of the centre note provides the best fit to the data in terms of minimising the pitch error. The temporal evolution of tonal reference during a piece revealed patterns of tonal drift in some singers, others appeared random, yet others showed periodic structure linked to the score. As a complement to errors of individual notes or intervals, we introduced a measure for the magnitude of drift, tonal reference deviation (TRD), and illustrated how it behaves using several examples.

Two types of factors influencing pitch error were investigated, those related to the singers and those related to the material being sung. In terms of singer factors, we found that pitch accuracy correlates with self-reported singing skill level, musical training, and frequency of practice. Larger intervals in the score led to larger errors, but only accounted for 2–3 cents per semitone of the mean absolute errors. On the other hand, the tritone interval accounted for 35 cents of error when it occurred, and in one case led to a large systematic error across many of the singers. We hypothesised that note duration might also have an effect on pitch accuracy, as singers make use of aural feedback to regulate their pitch, which results in less stable pitch at the beginnings of notes. This was indeed the case: a small but significant effect of duration was found for both the current note, and the nominal time taken from the onset of the previous note; longer durations led to greater accuracy. Many aspects of the data remain to be explored, such as the potential effects of scale degree, consonance, modulation, and rhythm.

6. ACKNOWLEDGEMENTS

Matthias Mauch is funded by a Royal Academy of Engineering Research Fellowship. Many thanks to all the participants who contributed their help during this project.

7. DATA AVAILABILITY

All audio recordings analysed here (and corresponding trajectory plots) can be obtained from `http://dx.doi.org/10.6084/m9.figshare.1482221`. The code and the data needed to reproduce our results (note annotations, questionnaire results, interruption details) are provided in an open repository at `https://code.soundsoftware.ac.uk/projects/dai2015analysis-resources`.

8. REFERENCES

[1] P. Alldahl. *Choral Intonation*. Gehrman, Stockholm, Sweden, 2006. p. 4.

[2] D.E. Brown. *Human Universals*. Temple University Press, Philadelphia, 1991.

[3] D. S. Crowther. *Key Choral Concepts: Teaching Techniques and Tools to Help Your Choir Sound Great*. Cedar Fort, 2003.

[4] S. Dalla Bella, J. Giguère, and I. Peretz. Singing proficiency in the general population. *The Journal of the Acoustical Society of America*, 121(2):1182, 2007.

[5] A. de Cheveigné and H. Kawahara. YIN, a fundamental frequency estimator for speech and music. *The Journal of the Acoustical Society of America*, 111(4):1917–1930, 2002.

[6] J. Devaney and D. P. W. Ellis. An empirical approach to studying intonation tendencies in polyphonic vocal performances. *Journal of Interdisciplinary Music Studies*, 2(1&2):141–156, 2008.

[7] J. Devaney, M. Mandel, and I. Fujinaga. A study of intonation in three-part singing using the automatic music performance analysis and comparison toolkit (AMPACT). In *13th International Society of Music Information Retrieval Conference*, pages 511–516, 2012.

[8] J. Fyk. Vocal pitch-matching ability in children as a function of sound duration. *Bulletin of the Council for Research in Music Education*, pages 76–89, 1985.

[9] C. M. Ganschow. Secondary school choral conductors' self-reported beliefs and behaviors related to fundamental choral elements and rehearsal approaches. *J. Music Teacher Education*, 20(10):1–10, 2013.

[10] D. M. Howard. Intonation drift in a capella soprano, alto, tenor, bass quartet singing with key modulation. *J. Voice*, 21(3):300–315, May 2007.

[11] H. Kawahara, J. Estill, and O. Fujimura. Aperiodicity extraction and control using mixed mode excitation and group delay manipulation for a high quality speech analysis, modification and synthesis system STRAIGHT. *Proceedings of MAVEBA*, pages 59–64, 2001.

[12] M. Kennedy. *The Concise Oxford Dictionary of Music*. Oxford University Press, Oxford, United Kingdom, 1980. p. 319.

[13] M. Mauch, C. Cannam, R. Bittner, G. Fazekas, J. Salamon, J. Bello, J. Dai, and S. Dixon. Computer-aided melody note transcription using the Tony software: Accuracy and efficiency. In *Proceedings of the First International Conference on Technologies for Music Notation and Representation (TENOR 2015)*, page to appear, 2015.

[14] M. Mauch and S. Dixon. pYIN: A fundamental frequency estimator using probabilistic threshold distributions. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2014)*, pages 659–663, 2014.

[15] M. Mauch, K. Frieler, and S. Dixon. Intonation in unaccompanied singing: Accuracy, drift, and a model of reference pitch memory. *Journal of the Acoustical Society of America*, 136(1):401–411, 2014.

[16] D. Müllensiefen, B. Gingras, and L. Stewart. Piloting a new measure of musicality: The Goldsmiths' Musical Sophistication Index. Technical report, Goldsmiths, University of London, 2011.

[17] M. Müller, P. Grosche, and F. Wiering. Automated analysis of performance variations in folk song recordings. In *Proceedings of the International Conference on Multimedia Information Retrieval*, pages 247–256, 2010.

[18] P. Q. Pfordresher and S. Brown. Poor-pitch singing in the absence of "tone deafness". *Music Perception*, 25(2):95–115, 2007.

[19] E. Prout. *Harmony: Its Theory and Practice*. Cambridge University Press, 2011.

[20] M. P. Ryynänen. Probabilistic modelling of note events in the transcription of monophonic melodies. Master's thesis, Tampere University of Technology, Finland, 2004. pp. 27–30.

[21] J. Salamon, E. Gómez, D. P. W. Ellis, and G. Richard. Melody extraction from polyphonic music signals: Approaches, applications, and challenges. *IEEE Signal Processing Magazine*, 31(2):118–134, 2014.

[22] R. Seaton, D. Pim, and D. Sharp. Pitch Drift in A Cappella Choral Singing. *Proc. Inst. Acoust. Ann. Spring Conf.*, 35(1):358–364, 2013.

[23] J. Swannell. *The Oxford Modern English Dictionary*. Oxford University Press, USA, 1992. p. 560.

[24] H. Terasawa. Pitch Drift in Choral Music, 2004. Music 221A final paper, URL `https://ccrma.stanford.edu/~hiroko/pitchdrift/paper221A.pdf`.