

Analysis of Intonation in Unison Choir Singing

Helena Cuesta¹, Emilia Gómez^{1,2}, Agustín Martorell¹, Felipe Loáiciga¹

¹*Music Technology Group, Universitat Pompeu Fabra, Barcelona (Spain)*

²*Joint Research Centre, European Commission*

helena.cuesta@upf.edu

Abstract

Choir singers synchronize their fundamental frequency (f_0) and timing when they perform together. In this work, we investigate several expressive characteristics of choir singing, with special emphasis on unison performances, to study how singers blend together and interact with each other by analyzing their f_0 dispersion, intonation, and vibrato. We use the spectral autocorrelation algorithm to obtain the f_0 trajectories of each individual singer and compute the f_0 dispersion as the standard deviation of the f_0 distribution. Using the Pearson correlation, we study the synchronization of singers in terms of their intonation and vibrato. Together with this study, we present a new choir singing dataset of voices singing together but recorded individually using directional close microphones. This dataset contains the individual recordings of 16 singers from a semi-professional 4-part choir (SATB) performing 3 different songs.

We found the f_0 dispersion in unison sections to be in the range of 16 to 30 cents, and the largest average f_0 dispersion was found in the basses. We observed that the intonation correlation varies a lot and is around 0.32 on average in all cases. The slowest piece was the one with the highest percentage of notes with vibrato, which had a frequency around 5 Hz for the whole dataset.

1. Introduction

Choir singing is probably the most widespread type of singing (Sundberg, 1987). The European Choral Association, for instance, represents more than 2,5 million singers, conductors, composers, and managers in over 40 European countries, reaching more than 37 million people in Europe active in the field of collective singing. Although these numbers suggest the relevance of choir singing research, there have been few research studies focusing on voice ensembles compared to the ones addressing solo singing.

Choir singing and solo singing differ in several acoustic characteristics such as loudness or formant frequencies. In a study about choir singing by Zadig et al. (2016), they state that the voices of a choir are commonly meant to blend together, which means that when they sing together, the audience is supposed to perceive one single voice, even though there are a greater number of them. A very relevant review on choir acoustics was published by Ternström (2003), who discussed several related studies. Some of these include an analysis of the intonation of intervals (Lottermoser and Meyer, 1960) or his own research on fundamental frequency (f_0) dispersion (Ternström, 1993).

Intonation is a complex phenomenon (Devaney, 2011) that describes how a pitch is sung in tune (Dai and Dixon, 2017). It is commonly referred to as the accuracy of pitch in singing or playing with respect to a specific tuning system; for example, the most typical in Western music is the equal tempered tuning system. Fundamental frequency (f_0)

dispersion is defined as the small deviations in f_0 between singers producing the same note at a unison performance. The agreement between all the voice sources is referred to as *degree of unison* (Sundberg, 1987): the larger the dispersion, the smaller the degree of unison. The investigation by Jers and Ternström (2005) is concerned with intonation, synchronization, and to what extent the singers of a voice section agree to each other. The authors carried out a multi-track recording of a 16-singer choir to analyze individual voices in terms of f_0 and found the dispersion to be between 25-30 cents¹. Another popular study on barbershop quartets (Hagerman and Sundberg, 1980) reported very small f_0 dispersion values: 4,3 to 16,9 cents. More recent studies include the analysis of interactive intonation in vocal ensembles (Dai and Dixon, 2017), where authors investigate how singers of an unaccompanied ensemble interact with each other in terms of intonation (i.e. pitch accuracy). Dai and Dixon designed a novel experiment and tested the intonation accuracy of individual singers by pitch and interval error. They found that singers interact with each other and that this interaction influences their intonation.

In addition to the intonation and f_0 dispersion, a very relevant aspect of choir singing is vibrato. In the early 30s, (Seashore, 1931) carried out one of the first studies on vibrato in singing using two approaches: a phono-photographic recording of musical performance and speech, and a psychophysical measurement on the perception of the vibrato produced synthetically by instruments. As a result of this study, he provided a definition of a vibrato as “*a periodic oscillation in pitch in which the extent of oscillation for the best singers averages approximately a half-tone and for string instruments approximately a quarter-tone*”. Seashore also claims that “*a good vibrato in music is a periodic pulsation, generally involving pitch, intensity, and timbre, which produces a pleasing flexibility, mellowness, and richness of tone*”. According to (Driedger et al., 2016), who proposed a template-based system for vibrato characterization, musicians extensively use vibrato as a musical effect to make their performance more expressive. Herrera and Bonada (1998) present a framework for the extraction and parameterization of vibrato and claim that long sustained notes become boring and uninteresting if their steady states have a strictly constant f_0 . This is why performers - and especially well-trained performers - tend to modulate both the frequency and the amplitude of the notes they produce, resulting in *vibrato* and *tremolo*, respectively. In the present study, we consider

¹ A *cent* is a logarithmic frequency unit. A musical semitone is subdivided into 100 cents (Driedger et al., 2016).

vibrato as characterized by its frequency or rate in Hertz and its extent or amplitude in cents. According to (Sundberg, 1987), an aesthetically reasonable range for the vibrato frequency is between 5.5 Hertz and 7.5 Hertz; this is why most methods to extract vibrato from audio signals usually restrict the search to this frequency range, approximately. Another relevant study about vibrato is the one by Prame (1992), who analyzed the performance of ten singers and found that the frequency of the vibrato typically increases at the end of the tones and that the average vibrato frequency across singers was 6.1 Hz.

The aim of our work is to study several expressive characteristics of choir singing, with special emphasis on unison performances, to study how singers blend together and interact with each other in terms of f_0 dispersion, intonation, and vibrato. We formulate three main hypotheses: (1) if we characterize unison performances by instantaneous mean f_0 and f_0 dispersion values, we can identify the perceived f_0 of the performance and quantify the degree of unison, respectively, (2) singers within the same choir section interact with each other in terms of intonation: they adapt their intonation to other singers' intonation, and (3) singers synchronize the frequency of their vibratos for the choir to be perceived as a single entity. We believe that obtaining information about these specific aspects of choir singing would be beneficial for the choir singing community, both for the singers in their practice and for the conductors to better understand their choir and design rehearsals accordingly.

2. Choral-singing Dataset Creation

A novel choral-singing dataset has been created for this study². We collaborated with the Anton Bruckner Choir³ from Barcelona (Spain) and organized a set of recording sessions in a professional studio with sixteen singers from the choir: four singers per section, i.e. four sopranos (S), four altos (A), four tenors (T) and four basses (B). We recorded the singers in groups, i.e. one session per section, singing in unison, with individual close microphones (with a cardioid polar pattern for directivity purposes) because we were interested in having separate tracks for each individual. In the first session, we recorded a video of the conductor of the choir while she conducted the performance. This video was displayed in the three remaining sessions and used for synchronization purposes. Singers also had the possibility to hear a piano reference (through headphones) for tuning purposes: *a cappella* singing often has the problem of f_0 drift during the performance, which happens when intonation moves away from the reference (Dai and Dixon, 2017). Since this phenomenon was not the focus of this study, we created a piano reference synchronized with the conductor movements.

Three pieces were selected based on the specific needs of the study, which were basically related to the language of the lyrics. These are the pieces we chose:

- *Locus Iste*, written by Anton Bruckner (Latin).
- *Niño Dios*, written by Francisco Guerrero (Spanish).
- *El Rossinyol*, popular Catalan song.

Overall, this dataset contains, for each of these pieces, the tracks of each individual singer (16 singers), together with the synchronized MIDI files of each choir section. Having the individual tracks also allows the user to create the unison mix for each section, as well as the whole choir performance. The dataset covers the frequency range between 87 Hz and 783 Hz, and is especially dense between 150 and 450 Hz. The notes have durations from 0.15 to 6.21 seconds, with an average of 0.84 seconds.

3. Methodology

Our methodology has three main parts: in the first one we study the characteristics of f_0 dispersion for the different choir sections; in the second one we analyze the correlation between the singers of the unison in terms of intonation, i.e. the degree of synchronization of their f_0 deviations; finally, in the last part of this research, we focus on describing vibrato in unison for singing, which is a very relevant aspect of singing.

Fundamental Frequency (f_0) Dispersion Analysis

The first part of this work aims at studying the f_0 dispersion found in unison choir singing. We obtained the f_0 trajectories for each singer individually using the spectral autocorrelation (SAC) method in (Villavicencio et al., 2015) as an f_0 estimation algorithm. We wanted to describe unison performances using a set of two descriptors, i.e. mean f_0 and f_0 dispersion, both computed from the individual f_0 trajectories of each singer.

The first step was to align audio recordings with their associated MIDI files. These MIDI files were time-synchronized with the audio recordings except for an offset at the beginning, which was manually corrected for all the cases. Then, we used SAC algorithm to obtain the f_0 predictions for each singer. The output of the f_0 estimation algorithm is a set of f_0 values computed frame-wise with a hop size of 5 ms: for each note of the score we have several values.

We divided the f_0 array into notes to study singer behavior per note. We were interested in the average dispersion within the note, following the study by Jers and Ternström (2005). To compute the note boundaries, we extracted the note onsets and offsets from the synchronized MIDI files using the Python library PrettyMidi (Raffel and Ellis, 2014) and segmented the f_0 array.

At this point, we had a set of f_0 values for each note of the score, for each piece and for each singer. We defined the f_0 dispersion as the standard deviation, σ_{f_0} , of the distribution of frequencies within a time frame. Besides the standard deviation, we also computed the mean f_0 (μ_{f_0}), which we hypothesize it corresponds to the pitch we perceive when we listen to the unison performance. We computed μ_{f_0} and σ_{f_0} for each note; however, in order to smooth the results, instead of computing the statistical metrics at each analysis frame (one value every 5 ms), we used a sliding window of 8 frames, therefore using 32 values at a time.

To obtain one single value per note for each metric we computed the average along the results for each window. Regarding the σ_{f_0} , we used a threshold before averaging the results: since the f_0 annotations were obtained with SAC and not manually, we had some errors in the f_0 trajectories that made the standard deviation reach very high values by mistake. In order to avoid this, we thresholded the values

² The dataset is published here: <https://doi.org/10.5281/zenodo.1286570>.

³ Anton Bruckner Choir: <http://edcantonbruckner.weebly.com>.

obtained at each window before computing the average. The threshold was set empirically to 60 cents by visualizing several examples and also taking into account the highest dispersion values obtained in other studies.

After the overall analysis of f_0 dispersion, we performed a finer analysis on some of the long notes - longer than 2 seconds, selected after a listening analysis of the pieces - of our dataset in order to study the trend of the f_0 dispersion in these cases. We expected this magnitude to be higher in the attack and release of the notes; however, we did not find any systematic pattern but only in some cases the dispersion is very high because of the attack imprecision, which would be an interesting topic for a study about the synchronization of singers in time.

Intonation Correlation

As mentioned above, singers produce slightly different frequencies even when singing the same note, i.e. f_0 dispersion. Since choir singing is about interacting with the other singers, it is likely that they adjust their intonation to others' intonation. We estimate here to what extent a singer is affected by how the other singers in the same section change their intonation, i.e. increasing or decreasing the produced f_0 . We expect these adjustments to be more or less linear: the goal of the intonation synchronization between singers is that the unison dispersion is as low as possible, which means that if a singer increases the produced f_0 , the singers around will probably adapt their intonations towards the same direction, either increasing or decreasing their f_0 . Assuming a linear correlation, we used the Pearson correlation coefficient, which according to (Papiotis et al., 2014), is the most common method used for quantifying linear dependence between two sets. This coefficient ranges between -1 (complete inverse linear correlation) and 1 (complete linear correlation), being 0 the lack of linear dependence between the two series.

Here, we were interested in studying if the general trend of the intonation was correlated between singers, not in the specific f_0 values. To capture this, we computed the derivatives of each of the f_0 trajectories, removing the continuous part of our data, and then segmented it into notes using the synchronized MIDI files. Instead of computing the Pearson correlation for each note, we also used small windows within each note in order to capture more precise information. When computing the correlation window-wise, we allowed for a small delay in the f_0 adjustments: if a singer modifies the intonation, it will take a while, i.e. a few samples, for the other singers to perceive this change and adapt accordingly. This is why we used a small window (20 frames, about 100 ms) around each sample. In Figure 1 we illustrate this effect with an example of one window: each line corresponds to the derivative of the f_0 trajectory for one singer. We see, for example, that alto 1 decreases around sample 7 and alto 2 imitates this behavior right after. To compute the Pearson correlation coefficient, we used the statistics module of Scipy, a scientific library for Python.

We computed the average of the coefficients of each window to obtain with a single correlation value for each note. We then analyzed the results manually to look for sections or pairs of singers that stand out from the others in terms of correlation. Notice that although coefficients can be either positive or negative, we are interested in their absolute value,

since two sets that show an inverse linear dependence might as well be correlated: if one voice decreases the intonation and another voice increases it to match the first one, we will get a negative correlation.

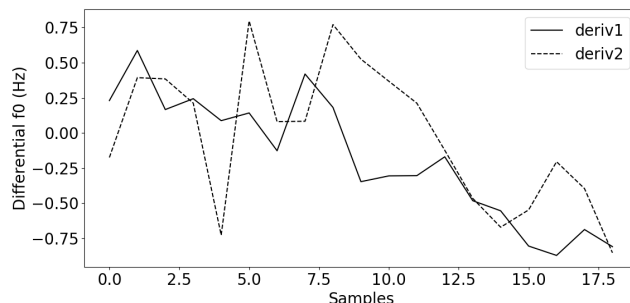


Figure 1. An example of the derivatives of the f_0 trajectories of two voices. This figure illustrates how one singer (deriv2) adjusts her intonation to another singer (deriv1). This window has a Pearson correlation coefficient of 0.57.

Vibrato Analysis

Vibrato is a very important aspect of singing, and most professional and semi-professional singers use it for expressivity purposes (Driedger et al., 2016). Seashore (1931) claims that it is present in the voices of all great artists in about 95% of their phonated time, even in transitions, attacks, and releases.

We used the vibrato extraction algorithm implemented in Essentia (Bogdanov et al., 2013) to find the segments of the performances that contain vibrato and to obtain the corresponding vibrato rates and amplitudes. The vibrato extraction algorithm from Essentia is based on the vocal vibrato detection from MELODIA (Salamon and Gómez, 2012), which is itself based on the method for vibrato characterization described by (Herrera and Bonada, 1998). Following the procedure from the previous stages, we segmented our data (i.e. f_0 trajectories) into notes using the score information, and then carried out a set of experiments described below.

Percentage of vibrato. We first measured the percentage of notes of the dataset that contained vibrato, sorted by singer and piece. Using the vibrato extractor described above, we checked the number of notes reported by the algorithm as having vibrato, and computed the percentage using the total amount of notes. This data is relevant to find out aspects such as whether a singer uses more vibrato than the others, or if singers are more likely to use vibrato in one of the pieces of our dataset.

Vibrato frequency and extent. Using the information extracted in the previous step about the notes that have vibrato, we then studied the vibrato parameters: frequency (or rate) and amplitude (or extent). We iterated through all the notes and kept only the ones with vibrato; notice that although a note is reported to contain vibrato, it does not mean that it is present all along the note. Instead, vibrato usually appears in different parts of the notes. Therefore, to study the frequency and amplitude of the vibrato we first located the segments where it was present, and then extracted both parameters given by the vibrato extractor. For each individual f_0 trajectory, we computed the average and standard deviation of vibrato frequencies and amplitudes.

Vibrato correlation. In previous steps, we studied the correlation between singers in terms of intonation. Following a very similar procedure, we studied the vibrato synchronization between singers of the same choir section (unison). We chose to use the vibrato frequency in this case; however, the vibrato extent could also be used to check whether singers adapt the parameters of the vibrato to the other singers. We computed the correlation between the vibrato produced by each pair of singers of the same section. We used the Pearson correlation coefficient as a metric to quantify, for each note, if the two singers produce similar vibrato rates, which is the same as in the study by Daffern (2017): she computes the correlation for specific relevant notes, i.e. the last note of the performance; however, we were more interested in the overall correlation values, so we averaged our results for each performance.

4. Results

In this section we present the results we obtained, organized following the same structure as the methodology: f_0 dispersion analysis, intonation correlation, and vibrato analysis.

F_0 Dispersion Analysis

Following the methodology described above, we extracted the f_0 trajectories and then computed the dispersion note-wise. The dispersion phenomenon is illustrated in Figure 2, where we display the f_0 trajectories corresponding to a single note of the soprano section. We observe that different and varying frequency values are produced by each singer, resulting in a set of similar yet different values of f_0 at each time instant. In Figure 3 we show the averaged results of the f_0 dispersion by choir section and by piece.

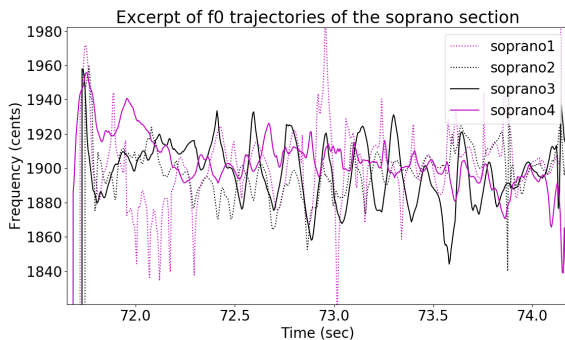


Figure 2. Excerpt of a unison performance by 4 sopranos singing a E5 (around 660 Hz) from *Locus Iste* (minute 1:12). Each line corresponds to the f_0 trajectory of an individual.

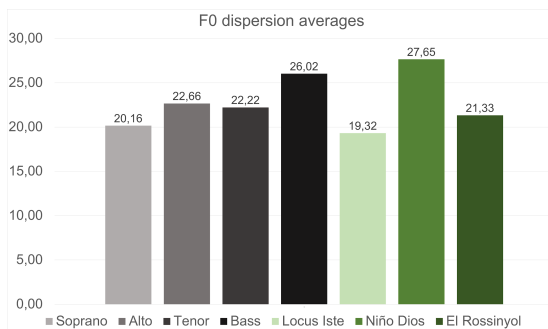


Figure 3. f_0 dispersion results averaged by section (grayscale) and by piece (greenscale).

By analyzing the individual results, we found the dispersion to vary between 16 and 30 cents, depending on the voice, which agrees with the results in (Ternström, 2002). In the Figure 3 we observe that although there are not very strong differences in dispersion between sections nor pieces, basses tend to show a higher dispersion, while sopranos obtained the lower values. Another relevant result is that the second piece was reported by the singers to be the most difficult to sing and has a higher average dispersion. This piece has a higher level of complexity in terms of rhythm, tempo, and intervals, so this result suggests that a more difficult piece might lead to higher dispersion values, therefore decreasing the degree of unison.

Intonation Correlation

Following the procedure detailed in the methodology, we obtained one correlation value for each analysis window, thus getting several correlation values for each note of the score. By manually inspecting some of the results, we realized we got significantly different correlation values between windows of the same note. This is illustrated in Figure 4, which includes a plot of the correlation results of all the analysis windows within a single note, i.e. one value for each window. We observe that the values oscillate, which suggests that the intonation adjustments are not systematic nor constant, and do not follow any specific pattern. However, we can see that there are highly correlated windows (very close to either 1 or -1) in these results, which also suggests that at some points, singers adjust their intonations.

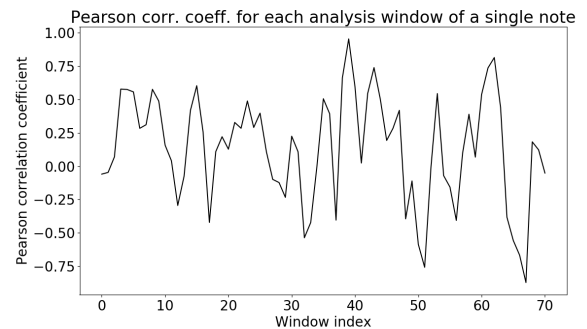


Figure 4. Evolution of Pearson correlation coefficient along one note.

Given the varying nature of the correlation in this context, it is difficult to generalize its trend. Our first approach was computing the average for each note and then manually study which pairs of singers are most correlated and which of the pieces had the highest correlation values. We did not average our results by singer or piece because some relevant information, such as maximum values, was lost. In Table 1 we summarize the results: for each piece and choir section, we report the maximum average correlation (max average) and also the maximum among the maximum correlation values (highest maxim.). For each of these two metrics we also show to which pair of singers they correspond (columns “pair of singers”) in order to check for repeated patterns, i.e. a pair of singers that has the highest correlation for all pieces.

Table 1. Summary of the results of the intonation correlation analysis.

Piece	Choir section	Max average	Pair of singers	Highest maxim.	Pair of singers
Locus Iste	Soprano	0.31	3-4	0.6	2-3
	Alto	0.33	3-4	0.7	1-4
	Tenor	0.36	1-2	0.58	1-2
	Bass	0.34	1-3	0.52	3-4
Niño Dios	Soprano	0.34	1-3	0.88	2-3
	Alto	0.33	3-4	0.78	1-4
	Tenor	0.33	2-3	0.82	1-2
	Bass	0.32	1-3	0.78	3-4
El Rossinyol	Soprano	0.33	1-4	0.66	2-3
	Alto	0.32	3-4	0.57	3-4
	Tenor	0.33	2-3	0.73	2-3
	Bass	0.32	2-3	0.58	2-3

On one hand, we observe in Table 1 that the average correlation is similar and around 0.32-0.33 in all cases. This particular result first suggests that there is not a specific pair of singers that is more synchronized than the rest. Additionally, it also shows that, on average, the correlation between singers in terms of intonation is quite low, even lower than 0.5. On the other hand, if we analyze the maximum correlation values (5th column in Table 1) we observe that if we look at the first two pieces (*Locus Iste* and *Niño Dios*), the pairs of singers with the highest maximum correlation are the same: sopranos 2 and 3, altos 1 and 4, tenors 1 and 2, and basses 3 and 4. In the last piece (*El Rossinyol*), however, this changes except for the sopranos. Although this might be a coincidence and this highest value is not representative of the whole data, it is likely that a deeper analysis of these singers would reveal more relevant patterns and more interesting information about the intonation interaction between singers.

Vibrato Analysis: Percentage of vibrato

Using the vibrato extractor described in previous sections we computed the percentage of notes that had vibrato for each piece and choir section. Notice that we computed the percentage for each singer, and then averaged the results for each section. The results are displayed in Figure 5 in the form

of a bar chart, where we observe that in general, singers produce vibrato in less than 50% of the notes. However, these results need to be analyzed taking into account the context: these singers are part of a semi-professional choir, and although most of them have some kind of singing training, in general they are not professional singers, which probably means that they do not have a complete control over their vibrato. In Figure 5, we also observe that the first piece (*Locus Iste*) has a significantly higher percentage of notes with vibrato, which might be a result of the piece the slowest tempo.

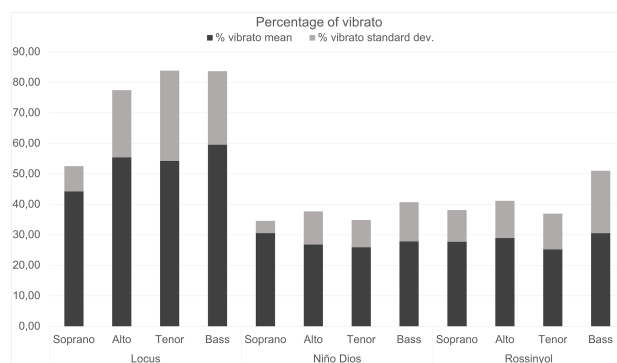


Figure 5. Percentage of notes with vibrato separated by piece and choir section.

Vibrato frequency and extent. In Figures 6 and 7 we display the results of our analysis of the frequency and amplitude of the vibrato, averaged and separated by section and piece. We observe that all vibrato frequencies are around 5 Hz, with quite low standard deviations, which is consistent with the results presented by Sundberg (1987) and Prame (1992). Regarding vibrato amplitude (Figure 7), we do not find any repeated pattern in the results, although we observe that for the second piece, in general, the amplitude of the vibrato is larger, especially for the basses. This is a surprising result, since this piece was reported to be the most difficult to sing by the singers, as well as the one with a fastest tempo, therefore we would expect the vibrato to be subtler or even missing.

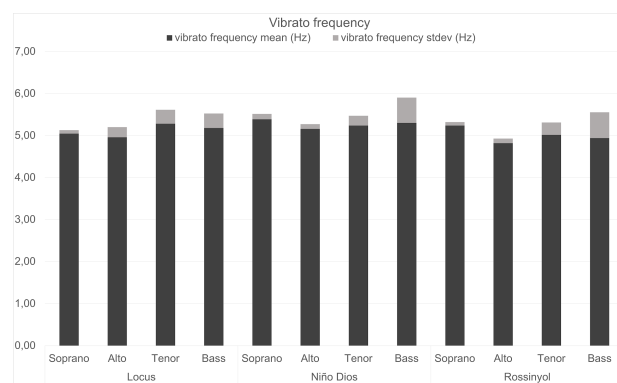


Figure 6. Vibrato rate/frequency averaged by choir section and piece. The values are displayed in Hz.

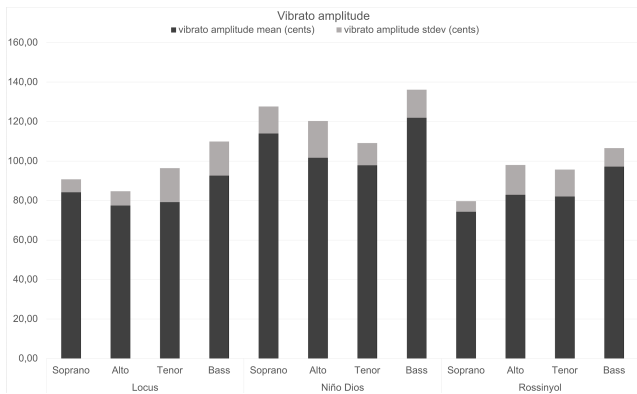


Figure 7. Vibrato amplitude/extent averaged by choir section and piece. The values are displayed in cents.

Vibrato correlation. For this part we present the pairs of singers that have an average vibrato correlation higher than 0.3. This threshold was chosen by sorting all the values (high to low) and keeping only those which differed less than 0.1 from their previous ones. The most relevant result is that the piece *Niño Dios* has the highest vibrato correlation: 0.56 for basses 4 and 2, and several values ranging from 0.36 to 0.47 for soprano pairs. However, these results are not very consistent nor explanatory, so we would need more data and more singers to corroborate and broaden them.

Conclusion

The main aim of this work was to study the synchronization and interaction between singers of a choir in terms of their f_0 dispersion, intonation, and vibrato. We characterized the unison performances by computing the mean f_0 and the f_0 dispersion and found the latter to lie in the range between 16 and 30 cents, being especially large in the bass section of the choir. We used the Pearson correlation coefficient to estimate how singers' intonation is affected by other singers when performing together, and our results show that the intonation correlation fluctuates a lot. Although we observed that a few pairs of singers had high correlation values in a few examples, the average was around 0.32 in all cases. With more data, we would be able to extract more relevant information. We have already planned a new recording session with a choir to extend our working datasets. We finally used a vibrato extractor to compute its presence in our dataset and compare it between different singers. The synchronization of the vibrato rate is a characteristic we would expect from more professional singers that have a much higher control of their voices.

Acknowledgements. This work is partially supported by the Spanish Ministry of Economy and Competitiveness under the CASAS project (TIN2015-70816-R) and the European Commission under the TROMPA project (H2020 770376). Many thanks to all the singers of the Anton Bruckner Choir who took part in this project.

References

Bogdanov, D., Wack, N., Gómez Gutiérrez, E., Gulati, S., Herrera Boyer, P., Mayor, O., ... & Serra, X. (2013). Essentia: An audio analysis library for music information retrieval. In *Proceedings of the 14th ISMIR Conference*.

- Daffern, H. (2017). Blend in singing ensemble performance: Vibrato production in a vocal quartet. *Journal of Voice*, 31(3), 385-e23.
- Dai, J., & Dixon, S. (2017). Analysis of Interactive Intonation in Unaccompanied SATB Ensembles. In *Proceedings of the 18th ISMIR Conference*.
- Devaney, J. (2011). An empirical study of the influence on musical context on intonation practices in solo singers and SATB ensembles. *Doctoral dissertation*. McGill University. Montreal, Canada.
- Driedger, J., Balke, S., Ewert, S., & Müller, M. (2016). Template-based vibrato analysis in music signals. In *Proceedings of the 16th ISMIR Conference*, p. 239-245.
- Hagerman, B., & Sundberg, J. (1980). Fundamental frequency adjustment in barbershop singing. In *TMHQPSR Speech, Music and Hearing: Quarterly Progress and Status Report*, 21(1), 28-42.
- Herrera, P., & Bonada, J. (1998). Vibrato extraction and parameterization in the spectral modeling synthesis framework. In *Proceedings of the Digital Audio Effects Workshop (DAFX98)*.
- Jers, H., & Ternström, S. (2005). Intonation analysis of a multi-channel choir recording. In *TMHQPSR Speech, Music and Hearing: Quarterly Progress and Status Report*, 47(1), 1-6. ISO 690
- Lottekmoser, W., & Meyer, F. (1960). Frequenzmessungen an gesungenen Akkorden. *Acta Acustica united with Acustica*, 10(3), 181-184.
- Papiois, P., Marchini, M., Perez-Carrillo, A., & Maestre, E. (2014). Measuring ensemble interdependence in a string quartet through analysis of multidimensional performance data. In *Frontiers in psychology*, 5, 963.
- Prame, E. (1994). Measurements of the vibrato rate of ten singers. *The journal of the Acoustical Society of America*, 96(4), 1979-1984.
- Raffel, C., & Ellis, Daniel P.W. (201) Intuitive analysis, creation and manipulation of MIDI data with pretty_midi. In *Proceedings of the 15th ISMIR Conference*. Late breaking demo paper.
- Salamon, J., & Gómez, E. (2012). Melody extraction from polyphonic music signals using pitch contour characteristics. In *IEEE Transactions on Audio, Speech, and Language Processing*, 20(6), 1759-1770.
- Seashore, C. E. (1931). The natural history of the vibrato. In *Proceedings of the National Academy of Sciences*, 17(12), 623-626.
- Sundberg, J. (1987). *The Science of the Singing Voice*. Northern Illinois University Press.
- Ternström, S. (1993). Perceptual evaluations of voice scatter in unison choir sounds. *Journal of Voice*, 7(2), 129-135.
- Ternström, S. (2003). Choir acoustics: an overview of scientific research published to date. *International Journal of Research in Choral Singing*, 1(1), 3-12.
- Villavicencio, F., Bonada, J., Yamagishi, J., & Pucher, M. (2015). Efficient Pitch Estimation on Natural Opera-Singing by a Spectral Correlation based Strategy. In *IPSJ Technical Report*.
- Zadig, S., Folkestad, G., & Lyberg Åhlander, V. (2016). Multi-track recordings of choral singers: Development and validation of a method to identify activities and interaction in the choral voice, based on recordings of the individual singers. *Bulletin of empirical music education research*, 7(1), 1-20.