Open Library of Humanities

Variation and change over time in British choral singing (1925–2019)

Edward J. Marshall*, University of Glasgow, UK, edward.marshall@glasgow.ac.uk

Jane Stuart-Smith, University of Glasgow, UK, jane.stuart-smith@glasgow.ac.uk

John Butt, University of Glasgow, UK, john.butt@glasgow.ac.uk

Timothy Dean, Royal Conservatoire of Scotland, UK, t.dean@rcs.ac.uk

*Corresponding author.

The front vowels of Received Pronunciation lowered in quality over the twentieth century (Wells, 1982b; Fabricius, 2007; Bjelaković, 2017). Connections between choral singing and Southern Standard British English (SSBE) have been made in musicological literature (e.g., Potter, 1998), and with specific reference to the Choir of King's College, Cambridge (Sagrans, 2016; Day, 2018). At King's and places like it, choristers are likely trained to sing using features of SSBE. This study investigates whether the SSBE pattern of front vowel lowering over time is also found in choral singing, both in a dialect area where SSBE is widely spoken (Cambridge) and in a non-SSBE dialect area (Glasgow). Two electronic, automatically segmented diachronic corpora were constructed in LaBB-CAT (Fromont & Hay, 2012). F1 and F2 were extracted for the vowels /i ɪ ɛ a ɑ/. Both datasets were analysed with Bayesian linear mixed models using `brms` in R. A main effect of lowering is supported by the data, priors and models. This paper provides evidence that British choral singing has changed over time and that this change follows a well-evidenced change in SSBE speech. We also find evidence for a shared non-regional choral accent emerging in terms of front vowel height.

1. Introduction

This paper presents an account of variation and change in the front vowels of British choral singing. Previous sociophonetic research on singing has focused on solo singing in popular styles using qualitative and quantitative variationist methods (e.g., Trudgill, 1983/1997; Beal, 2009; Krause & Smith, 2017; Caillol & Ferragne, 2019). Previous acoustic phonetic research on singing has focused on the acoustics of pitch and tuning, and the "singer's formant" (Sundberg, 1987), amongst other things. Musicological research has discussed choral "sound" in terms of tone colour, for example, of King's "pure dispassionate quality" (Day, 2018). The study presented here is part of an ongoing interdisciplinary program of research whose ultimate goal is to identify what constitutes choral "accent" (from a linguistic perspective) and choral "sound" (from a musicological and performance/practice perspective), exploiting theoretical perspectives, constructs and methods from sociophonetics, musicology and singing performance and practice, together in Marshall (2023a), *AHRC project code:* 2284740.

Here we present the results of an analysis of a new real-time corpus designed to interrogate choral singing (1925–2019) in two classical choirs from two different English dialect areas, in England and Scotland. The Choir of King's College, Cambridge, is located in the south of England, in a Southern Standard British English (SSBE, formerly Received Pronunciation) dialect area (Hughes, Trudgill, & Watt, 2012). The Orpheus and its continuation, the Phoenix, choirs are located in Glasgow, in a Central Scots/Scottish English dialect area (Stuart-Smith, 1999). We focus on acoustic vowel quality in the English vowels /i ɪ ɛ a ɑ/, which, following the convention of recent British dialectology, we refer to as the FLEECE, KIT, DRESS, TRAP, and BATH vowels (Wells, 1982a). Our results provide the first evidence to support the notion of a "British choral accent," close to spoken SSBE in terms of phonemic inventory and phonetic realisation, which is used not only in Cambridge, as we might expect, but also for Scottish choral singing. We also find that there is diachronic variability in the front vowel system in both choirs, which on the one hand reflects documented variation and change over time in spoken SSBE (Harrington, Palethorpe, & Watson, 2000), and on the other, aligns with the performative interventions of particular choral directors which have been already observed in the musicological literature (Day, 2018).

Choral singing is a conservative tradition. Some of the regular repertoire has been sung for at least a thousand years. In a similar vein to Leech-Wilkinson (2009), if we had recordings of singing in English from the past 500 years we would be able to trace how pronunciation has changed over the entire evolution of Modern English. Whilst we will never have those recordings, we do have data for the past 100 years. For these data, there is both innovation in repertoire and also repetition of historical repertoire. This makes historical choral recordings ideal for investigating musical and linguistic–phonological variation over time.

The paper is structured as follows: First, in Section 1, we contextualise our study with respect to the previous work on phonological variation in singing, acoustic phonetic research on the singing voice, and musicological research on choral sound, and how this is achieved by choral directors. In Section 2, we introduce the choirs (the Choir of King's College, Cambridge, and the Glasgow Orpheus and Phoenix choirs), and our methods for analysing formants in (unaccompanied) choral singing, as well as outline the Bayesian statistical modelling strategy adopted here. We then present the results for each choir separately and together in Section 4. We conclude by discussing our findings with respect to the research questions and previous motivating research in Section 5.

1.1 Previous sociophonetic research on singing

The first sociophonetic study on singing was carried out by Trudgill (1983/1997), looking at British pop singers performing in the 1960s–70s. The groups he considered, such as the Beatles and the Rolling Stones, showed American accent features when they were singing compared to their spoken (British) pronunciation (see also, Simpson, 1999). Trudgill's work was followed by a series of studies, including Beal (2009) on the Sheffield rock band Arctic Monkeys, Krause and Smith (2017) on Glasgow indie bands the Twilight Sad and the Unwinding Hours, Yang (2018), on Lenka, a pop singer from Australia, and Caillol and Ferragne (2019), on the British heavy metal bands Def Leppard and Iron Maiden. These studies used a combination of a variationist Labovian approach correlating pronunciation variability with artists over time (e.g., Labov, 1972) combined with qualitative information. Data in this work is usually the result of auditory coding, although there are some small-scale quantitative analyses of acoustic data for popular solo singing (Gibson & Bell, 2012), and Yang (2018) provides qualitative acoustic analysis in the form of spectrographic analysis of some vowel qualities.

Early sociolinguistic studies of singing tended to neglect the impact of the aesthetic requirements of singing (Morrissey, 2008; Gibson & Bell, 2012). Morrissey notes the importance of the sonority of individual sung speech sounds; the more sonorous a sound, the better it "carries" the tune. Another important contribution that Morrissey makes is a theoretical construct for considering singing style, specifically the notion of the "reference style." He regards the "mid-Atlantic" pseudo-American accent and SSBE (formerly RP) as "dominant reference style[s]," and notes how deviating from these reference styles can be marked, such that popular singers can use deviation from the reference style as an effect in their performances. Gibson (2019) develops these ideas further, introducing the idea of Standard Popular Music Singing Style (SPMSS).

Classical styles of singing within the overarching classical "reference style" (Morrissey, 2008) have received little attention within sociophonetics. Wilson investigates the language ideologies at play in choral rehearsals in Trinidad (Wilson, 2014; 2017). Wilson (2014) is

the first sociolinguistic study to engage with a classical singing style. Wilson finds that many participants associate ideal forms of choral pronunciation with British English features, arguing that, following Gibson and Bell (2012)'s view on American accents in popular singing, "British English has become institutionalised with regard to choral singing – it is so consistently associated with choral singing that it has begun to function as the default style of this activity" (Wilson, 2014, p. 316).

Acoustic work on classical singing has almost always focused on operatic technique, despite the fact that choral singing is probably the most widespread type of singing (Sundberg, 1987). Hence our study focuses on the conservative genre of Western classical choral singing, exploiting the fact that, with recordings from over a century, choral singing is an untapped resource for researching variation and change over time from both musical and linguistic perspectives. The following section will present singing and choral leadership perspectives on pronunciation in classical singing.

1.2 Singing perspectives on pronunciation in classical choral singing

In many UK cities, including Glasgow and Cambridge, there is a long history of classical choral singing, with recordings dating back to the 1920s (e.g., the Glasgow Orpheus Choir, or the Choir of King's College, Cambridge). These choirs often sing in Latin or modern European languages (and languages from further afield), however, this study focuses on pronunciation of choral singing in the English language. Musicologists acknowledge that choirs typically have a "sound," which includes how choirs produce speech sounds while they sing (Sagrans, 2016): a choral "accent." But neither phoneticians, singers, nor choir directors have a clear understanding of how such choral sound-accents are achieved, how they arise, and/or are maintained.

Manuals exist for classical solo singers which give pronunciation advice, with explicit appropriate norms for singing in other languages (e.g., Adams, 2008; A. Johnston, 2016). For example, Adams provides guidance for pronunciation when singing in French, Italian, German – but only talks about the pronunciation of English when it causes issues for singing in the foreign language, e.g. "in English, vowels in unstressed positions almost always neutralise to /ə/. English speakers unwittingly carry this habit over into Italian" (Adams, 2008, p. 6).

There are also manuals for choir directors (e.g., Marvin, 1991; Crowther, 2003; Burns & Kydd, 2013; Hollins & Vango, 2022), however, there is generally less focus in the literature on how to develop a choir's sound, or what a choir *should* sound like. A few exceptions such as Hollins and Vango (2022) provide warmups for choirs and explain what each exercise is meant to achieve (sonically, or for the singers' vocal wellbeing), but these tend to focus more on musical rather than linguistic aspects. There is practical advice for choir leaders new to running or setting up groups (e.g., Burns & Kydd, 2013), though this resource tends to focus more on

the operational side of choir leading, including setting rates, advertising performances, and the practicalities of running sessions.

Developing a unified vowel quality and pitch has been identified by practitioners as of central importance for choral sound (e.g., Powell 1991; Marvin, 1991; Crowther, 2003; Hollins & Vango, 2022). A certain degree of homogeneity is a necessary precursor for the development of a recognisable style (Sagrans, 2016). In the next section we will explore musicological findings about an English choral style.

1.3 Musicological perspectives on singing pronunciation

The musicological literature supports the notion of an English style of classical choral singing. Sagrans writes that the "[King's sound] is similar to a broader 'English sound' found among other Oxbridge college choirs…" (Sagrans, 2016, p. vii). In *I Saw Eternity the Other Night*, Day (2018) explores the development of "English Cathedral Sound" particularly focusing on the Choir of King's College, Cambridge and its commercial recording success. While Day mentions the particular style of enunciation which developed as the English choral tradition developed, he largely ignores accent features. There is a frequent discussion of musical and timbral aspects including the renowned "pure, dispassionate quality" of the sound cultivated from English treble voices (Day, 2018, p. 5). Potter describes the style of King's under the direction of Boris Ord and David Willcocks as "exaggerating consonants and vowels far beyond the demands of clarity" (Potter, 1998, p. 117). Both Day and Potter suggest that due to the fame and commercial success that King's choir propagated, their particular style and pronunciation norms spread to church choirs and choral societies throughout the UK. Potter continues that "the pronunciation established criteria by which excellence could be measured, and the singing was judged in part by the excellence of its own pronunciation rather than by the success or otherwise of strategies to put across the meaning" (Potter, 1998, p. 117). Sagrans describes the "unvarying nature of the King's sound" (2016, p. 65), whereas Day supports a shift in the style of the King's sound changing over time, writing:

> Willcocks certainly thought that a choir had a vocal identity, like a singer, and that it would be "fussy and pernickety" to expect a choir to vary the tone-quality it produced. Nevertheless, even under his disciplined control, recordings confirm that the style evolved (Day, 2018, p. 173).

Thus far, we have shown how developing a unified sound is important to choral directors, and that musicologists of choral singing have identified an "English" style of choral singing. This style is particularly tied to the Choir of King's College, Cambridge – and there may be evidence of the King's style evolving over time. The following section will explore evidence of regional variation in singing.

1.4 Classical choral singing in different regional dialect areas

Choralists, musicologists, and the singing literature more broadly, show some awareness of the issue of regional variation, as shown in the following extracts. In relation to singing pronunciation more generally, Adams notes that:

> English has closed /u/ as in *boot* and open /ʊ/ as in *foot*. Italian has only the closed position for the vowel-letter *u*. The sound /u/ is subject to great regional variation in English-speaking countries. This fact can present problems when this vowel sound is sung (Adams, 2008, p. 6).

Regarding choral singing specifically, Marvin writes that:

> In warmups concentrate on pure vowels: oo, oh, ah, eh, ee; do not allow diphthongs or regional accents to violate the purity of vowel color. Good intonation cannot be achieved until the vowels within and throughout all sections are unified (Marvin, 1991, p. 31).

As we have seen, choral literature mostly focuses on creating a homogeneous sound, though the question remains, whose sound, or which sound is being aimed for? It would appear from Marvin (1991) that the unified vowel sound being aimed for may be a non-regional one. This insight matches thinking in musicology, as Potter (1998) suggests a connection between the non-regional accent of British English, Received Pronunciation, and singing, writing:

> No research appears to have been done on the relationship between RP [Received Pronunciation] and singing and it is not yet possible to establish a particular link between them, but the ideological force of associating singing with high-status pronunciation in England can only have made classical singing more elitist (Potter, 1998, p. 65).

This leads us to ask, if it might exist, where does a classical choir accent come from? There are many factors that may influence choral accents from social, linguistic, and musical domains. These include individual singers' accents, wider regional accent(s), choir director's accent, genre/aesthetic considerations, and even the sonority of the individual speech sound. What are the accepted norms of choral singing? Are there shared norms between choirs of different dialect areas – i.e., a shared phonology with different realisations? As Potter suggests above, it is possible that British classical choral singing is based on the phonology of SSBE. Wilson (2014) argues that Western classical choral singing has its own phonology based largely on singing technique. These insights brings us back the notion of the reference style put forward by Morrissey (2008), but for a classical singing genre (as opposed to the popular genres previously investigated). Alternatively, the sung phonology may be influenced by local linguistic varieties.

There is also the element of diachronic change to consider. It is well-documented that the front vowels of Received Pronunciation (now SSBE) lowered in the twentieth century (e.g., Wells,

1982a; Fabricius, 2007; Bjelaković, 2017). As we have seen, the vowel quality of British choral singing is thought by musicologists to be based on SSBE pronunciation (e.g., Potter, 1998) and also with specific reference to the Choir of King's College, Cambridge (Day, 2018). At King's, and similar institutions, choristers may speak a variety of SSBE, but it is likely they are encouraged to sing with SSBE accent features by their choir directors. This study investigates whether the SSBE pattern of front vowel lowering over time is also found in choral singing, in a dialect area where SSBE is widely spoken (Cambridge) and in a non-SSBE dialect area (Glasgow). We will now outline previous acoustic work on singing.

1.5 Acoustic analysis of singing

In speech, vowels are transient phenomena often highly affected by coarticulation (coloured by the adjacent vowel or consonant sounds). For example, this is particularly noticeable for rounded back vowels in alveolar environments e.g. *toot* /tuːt/ (Hillenbrand, Clark, & Nearey, 2001). In singing, the vowel is often held for longer periods of time, usually determined by the musical rhythm or setting. Consequently, singers must be able to keep the shape of the vocal tract constant for long periods of time (Gregg & Scherer, 2005). In addition, the "homogenous timbre" required in singing is produced by manipulations of the vocal tract that are not typical of the vowels of speech (Deme, 2014). Classical singing training has been shown to have a general centralising effect on vowels (Dromey, Heaton, & Hopkin, 2011). The phenomenon of vowel "centralisation" has been noted particularly in sopranos singing in the high range (Hollien, Mendes-Schwartz, & Nielsen, 2000). Above 500 Hz, vowel identification is biased towards /a/ and above 700 Hz, vowel identification reaches chance performance (Friedrichs, Maurer, Suter, & Dellwo, 2015). However, it has been noted that this may have to do with the style of singing (Smith & Scott, 1980) and that it is mainly vowels "sung with classical vocal training" that have a tendency for "migration" (Gregg & Scherer, 2005, p. 209).

Previous acoustic work on classical singing and choral singing is usually framed as an investigation of vocal production, that is, how the sound is produced, rather than the qualities of the resulting sound. Choral sound relies on the singers' ability to match pitch, loudness and vowel quality. There has been much work on pitch matching, as it is relatively straightforward to examine (e.g., Jones, 1989; Demorest & Clements, 2007; Riegle & Gerrity, 2011; Lévêque, Giovanni, & Schön, 2012; Shekar & Fujioka, 2014). We know that singers match small pitch manipulations and they can do so quickly in real time (Grell, Sundberg, Ternström, Ptok, & Altenmüller, 2009). There has also been acoustic work on ideal signal-to-noise ratio in choirs and how the arrangement of individuals can improve a choir's sound by helping everyone hear themselves in the mixture (Sundberg & Ternström, 1986). There has even been work on singers' ability to match rate of vibrato (King & Horii, 1993); however, there is little to no research on vowel quality matching, due to the complexity and multidimensional nature of vowel quality (timbre).

Both phoneticians and singers are interested in formants, but in different ways. For phoneticians, especially F1 and F2 are important cues to vowel quality. Work on formants in singing has focused on what is known as the "singer's formant" in the F4 region, which gives classically-trained male singers the "pingy" quality of their sound and allows them to be heard over a large orchestra (Sundberg, 1974). However, there has been little research on F1 and F2 from a sociophonetic perspective on acoustic vowel quality in singing of any genre. Yang (2018) gives a qualitative comparison of spectrograms to inform their auditory coding of variants employed by the singer Lenka. Caillol and Ferragne (2019) give a small-scale comparison of vowel formants of FOOT /ʊ/ and STRUT /ʌ/ in the singing and speech of the lead singers of the heavy metal bands Def Leppard and Iron Maiden. They find that, while the singers have no FOOT–STRUT split in their speech, they do in their singing, and this reflects an adaptation towards the (standard) British or USA-5 models both of which have the FOOT–STRUT split (Caillol & Ferragne, 2019). To date little is known about how formants behave in choral singing. Sundberg (1987) asked:

> What happens to these formant frequency differences in a good choir? Do the members compromise so that all choir members agree on approximately the same formant frequencies, or do all section colleagues arrive at such an agreement, or is this an unimportant factor? In any event, the formant frequency distribution within a choir must be relevant to the choral timbre. Perhaps there are differences in this regard between choirs and perhaps also between the choral traditions in various countries (Sundberg, 1987, p. 145).

These questions raised by Sundberg (1987) remain of central importance to an understanding of choral sound/accent, and, as far as the authors are aware, are still yet to be answered. Particularly pertinent to this study is the final suggestion which we will address: whether there are differences in the formants produced by choirs from different dialect areas. We will now outline the research questions that guide this study.

1.6 Research Questions

This study investigates the vowel quality of choral singing using sociophonetic acoustic methods and specifically in two dialect areas: Cambridge and Glasgow. Within the limitations of a single paper, we restrict our scope to the front vowel system, i.e., FLEECE, KIT, DRESS, and TRAP, and the vowel pair TRAP and BATH. The primary research questions of this research are:

1. What is British choral singing like in terms of vowel phonology?
    (a) Specifically, what is the inventory and realisation of the front vowels in British choral singing?

     (b)   Is the front vowel phonology shared across Cambridge and Glasgow, i.e., is there a common British choral accent?

  2. Is there evidence of variation and change in British choral singing over time?

     (a)   Does British choral singing show the same pattern of front vowel lowering over time, as exhibited in spoken Standard Southern British English?

     (b)   Does this change occur equally in a dialect area where General British English is widely spoken (Cambridge) and in an area where it is not (Glasgow)?

Our predictions are as follows:

If the King's choral accent is based on SSBE, then we would expect the vowels in King's to reflect SSBE phonology. If there is a common British choral accent, we would also expect the vowels of the Glasgow choirs to show SSBE phonology, despite the Glasgow choirs being located in a Scottish dialect area. Furthermore, we might expect any diachronic change in both choirs to reflect the vowel lowering, particularly for TRAP, that has been observed for RP/SSBE over the twentieth century (Wells, 1982a; Fabricius, 2007; Bjelaković, 2017).

In SSBE, the vowel lexical sets TRAP and BATH select /a/ and /ɑː/ respectively, whilst in Scottish Standard English, both sets take /a/. Given that Scottish Standard English shows a single vowel phoneme /a/ for both TRAP–BATH sets – its realisation is typically an unrounded open central vowel, [a̠]. We predict that the SSBE TRAP–BATH split will be present in both corpora. If recordings from Glasgow present separate TRAP and BATH phonemes and front vowel lowering over time, then we have preliminary evidence of a supralocal norm for British choral singing.

2. Methodology

In order to answer these research questions, two electronic time-aligned corpora were constructed in LaBB-CAT (Fromont, 2019). The Glasgow corpus consists of commercially released recordings of the Glasgow Orpheus (1906–1951) and Glasgow Phoenix (1951–present) choirs with audio recordings from 1925 to the present day. The King's corpus consists of commercially released recordings and public broadcasts of the Choir of King's College, Cambridge, with audio recordings from 1945 to 2019.

2.1 Ethics

As we are working with commercially released and/or public broadcasts, this study did not require ethical approval. The acoustic data extracted from the recordings are available on the OSF https://osf.io/3vjxm/. If you would like access to the corpus for checking, please contact the corresponding author.

2.2 Sample

2.2.1 The Choir of King's College, Cambridge

King's was selected for its status as the prototypical collegiate choir. Sagrans argues that "the King's sound is a particularly high-profile example of a broader 'English sound' for choral performance" (Sagrans, 2016, p. 29). Musicologists have remarked that the development of the King's style and their frequent broadcasts led to a shift in choral singing practices in church choirs across the country (Potter, 1998; Day, 2018). Sagrans (2016) describes how the King's sound is similar to a wider English sound found in British early music vocal ensembles. This similarity is attributed to King's early recorded output containing a large proportion of early music, as well as the fact that many of the members of these new early music ensembles came from the universities of Oxford and Cambridge. If there were a place where a standard spoken accent was contributing to changes in choral singing style, it would be found at King's. The choir is formed of a combination of boy choristers and adult male choral scholars and lay clerks.

There is one recording made under Harold Darke during Boris Ord's service in the RAF from 1941–1945 which is included under Ord.[1] As there was only one album recorded by King's under Daniel Hyde at that point, the record was collapsed into the factor level SC = Stephen Cleobury. As shown in **Table 1**, for modelling purposes, time is coded as a four level factor for King's: Boris Ord = BO 1945–1958 ; David Willcocks = DW 1959–1974 ; Philip Ledger = PL 1976–1982 ; and, Stephen Cleobury = SC 1984–2019.

Corpus	Dates active	Director	Coding
King's	1929–1958	Boris Ord	BO 1945–1958
	1957–1974	David Willcocks	DW 1959–1974
	1974–1982	Philip Ledger	PL 1976–1982
	1982–2019	Stephen Cleobury	SC 1984–2019
	2019–present	Daniel Hyde	

Table 1: Choir of King's College, Cambridge time periods and choir directors.

2.2.2 Glasgow: Orpheus and Phoenix choirs

The Glasgow Orpheus and Phoenix choirs were selected in order to investigate an elite (auditioned) regional choir with a phonology that may differ from King's. The Orpheus and Phoenix choirs

---

[1] There is one record made under Arthur H. Mann from 1929 which is unaccompanied, however this recording was excluded as there were few recordings from the time and the recording is quite noisy. You can listen to a digitisation of the recording on YouTube here: (https://www.youtube.com/watch?v=iue8jGxlqVM).

have a long history of recordings and broadcasts. Both choirs are adult mixed-gender choirs. The Glasgow Orpheus Choir was conducted by Hugh S. Roberton, who led the choir from its inception. There has been little academic attention paid to the Orpheus, though there was some interest during its lifetime with an article in the *Musical Times* in 1925. Grace (1925) discusses the egalitarian principles that the Orpheus choir was based upon, including the annual vocal "examinations," the fact that the members did not pay to participate in the choir – but that the choir was entirely funded by ticket sales. Additionally, the article mentions the importance of fellowship and choral evangelism – the impact the choir had on the choral culture in Scotland as a whole – with alumni of the choir going on to conduct many large choruses or choirs across the country. The Orpheus ceased to exist in 1951 when Roberton stood down as director due to ill health and the choir unanimously agreed to disband (Roberton & Roberton, 1963). However, a large number of singers wanted to continue singing, so they founded the Glasgow Phoenix Choir in the same year (Glasgow Phoenix Choir, n.d.).

There are recordings of the Glasgow Phoenix Choir almost every other year since 1959 and to the present day. The first author collected all available recordings of the Glasgow Orpheus Choir (1925–1951) and the Glasgow Phoenix Choir (1959–present) that were produced in the UK. In addition to their own innovations, the Glasgow Phoenix Choir recorded the most popular repertoire of the Orpheus many times over the last 60 years. For example, there are nine recordings of "All in the April evening" (composed by Roberton) ranging from one recording conducted by Roberton himself, to a recording celebrating the centenary of the founding of the Orpheus choir in 2001.

Corpus	Dates active	Director	Coding
Glasgow	1901–1951	Hugh S. Roberton	HSR 1925–1951
	1955–1983	Peter Mooney	PM 1959–1975
	1983–1990	Peter S. Shand	MJS 1987–2016
	1991–2016	Marilyn J. Smith	
	2018–Present	Cameron Murdoch	

Table 2: Glasgow Orpheus and Phoenix choir time periods and choir directors

As shown in **Table 2**, due to the number of tokens available and their uneven distribution, for modelling purposes the Glasgow data were reflected in three factor levels: Hugh S. Roberton = HSR 1925–1951; Peter Mooney = PM 1959–1975; and, Marilyn J. Smith = MJS 1987–2016 (the latest time period containing the recordings made under Peter Shand and Cameron Murdoch). Hugh S. Roberton and Peter Mooney spoke with broadly Standard Scottish English accents, whereas Marilyn J. Smith has a Southern Standard British English accent.

The difference in the age and gender composition of the choirs analysed here could be considered a confound in the comparison of King's and Glasgow. However, we suggest an alternative perspective. Each choir is comprised of a mixture of vocal tract lengths; the Glasgow choirs both contain adult male and female singers, whereas King's is comprised of adult males and boy trebles. Here, we investigate large-scale general patterns of phonetic and phonological change in choral singing that relate to dialect differences. In both choral corpora, these patterns are present irrespective of the age/gender difference.

2.3 Recordings

To remove any effect of language, the recordings selected for analysis were restricted to English. We would expect choirs to sing differently when singing in another language e.g., French, German, and perhaps Latin in a UK context (Adams, 2008; A. Johnston, 2016). Scottish vernacular items (Scots) were included in the analysis and coded as Scottish Vernacular in the factor Genre for the Glasgow choirs.

For inclusion in this study, the recordings had to be unaccompanied, that is, not accompanied by any musical instruments, as this could affect the accuracy of the forced aligner for segmentation and also the reliability of the extracted acoustic data. The recordings had to be homorhythmic. That is, there is a simple texture with only one word being sung at a time, such as with hymn or psalm singing.[2]

The Glasgow corpus includes extracts from 178 tracks (songs) from 28 albums. The King's corpus includes extracts from 317 tracks from 50 albums. A full discography of the corpora, with Genre coding and more information about the soundfiles can be found on the OSF (https://osf.io/3vjxm/).

2.4 Segmentation using forced alignment

The text (lyrics) sung on a given track were found online or transcribed by ear. Single interval tiers were created in Praat (Boersma & Weeninck, 2018) and manually divided into chunks of approximately 10 seconds with a short silence between each chunk. The sound files were converted to mono and downsampled to 11,025 Hz in Praat. These recordings, along with the textgrids, were then uploaded to LaBB-CAT corpora annotation store (Fromont & Hay, 2012). Phonemes were annotated using the CELEX English dictionary (Baayen, Piepenbrock, & Gulikers, 1995). Word and phoneme boundaries were automatically aligned in LaBB-CAT using the Hidden Markov Model Toolkit (Young et al., 2009).

As far as the authors are aware, no aligner has been trained to work with singing, and certainly not choral singing. As anticipated, applying the HTK model to choral singing data

---

[2] In polyphonic music, there can be more than one word being sung at a time such that it would be impossible to align and to extract the formant data without it being a mixture of different vowel qualities and/or consonants.

had mixed results. Forced alignment provided a much better starting point for hand-correction than hand-segmenting the whole corpus from scratch. Hand-correction of alignment focused on stressed vowels. All word and segment tiers were hand-corrected by listening to the sound file and visually inspecting the alignment with the waveform and spectrogram.

2.5 Hand-correction of vowel phone labels

CELEX-English uses conservative Received Pronunciation canonical pronunciation, so it does not accurately map onto a Scottish Standard English phonology. However, labelling consistently allowed for direct comparison of the two (possible) varieties. For example, CELEX-English and RP (and SSBE) have two separate vowel phonemes for the low vowels TRAP–BATH, whereas in Scottish English there is typically one target phoneme, CAT. We can use this inaccurate characterisation of the Scottish variety to compare whether a hypothetically Scottish TRAP is different to an SSBE TRAP. We would expect the Scottish TRAP realisation to be more central than the SSBE realisation.

Our analysis focused on stressed vowels from the aligned corpus for each choir. Not only were boundaries manually corrected, but also vowel phone labels also needed to be corrected. Occasionally, the phone label from the phoneme dictionary was incorrect. For example, orthographic *<a>* in English is variably produced as the diphthong /eɪ/ as in FACE, or as /ə/ (schwa). However, *<a>* was categorically transcribed as /eɪ/ even when it is frequently the short, unstressed, mid central vowel /ə/ (schwa). These instances were manually corrected when they were found. This confusion is also often the case with the TRAP phoneme, which is often a mislabelled /ə/ vowel. Likewise, *<the>*, which is variably realised as FLEECE /ði/ or schwa /ðə/ is often mislabelled. Other less frequent words can also have multiple pronunciations, e.g., *<aye>* can be produced using either the FACE /eɪ/ or PRICE /aɪ/ phoneme. These instances were corrected as systematically as possible throughout.

Setting a spoken text to music, often entails lengthening of vowels which would be short and unstressed in speech. We therefore listened to every expected instance of unstressed vowels and assigned each instance to the respective full vowel quality according to how it was actually realised. For example, in *lady* the final vowel was usually sung as a full vowel [i], as opposed to a reduced vowel [ɪ] in speech. Even in the earliest recordings from King's, the word-final HAPP-Y vowel is almost categorically produced as /i/ as in FLEECE, rather than /ɪ/ as in KIT. For example, *<lady>* is produced /leɪdi/ rather than /leɪdɪ/, *<heavenly>* is produced /hɛvənli/ rather than /hɛvənlɪ/.[3] Following Gibson and Bell (2012), function words were included where they were unreduced, and all repetitions were included in the corpus.

---

[3] An anonymous reviewer points out that HAPP-Y being produced as FLEECE rather than KIT is an example of the early recordings not conforming to Received Pronunciation, most likely driven by musical-aesthetic demands.

2.6 Automatic formant extraction

F1 and F2 were extracted from 7 time points equally distributed between 20% and 80% of the way through the vowel interval for FLEECE, KIT, DRESS, TRAP, and BATH. Averaging over 7 time points should reduce coarticulatory effects as well as the coordination issue that comes with joint speech and singing. Formants were extracted in LaBB-CAT using the Praat function "Sound: To Formant (burg)" with default settings, and a formant ceiling of 5,000 Hz. Since we are approximating formant values from a mixture of vocal tract lengths, it is unknown what value is most suitable for choral singing (and joint speech). As we are mostly interested in F1/F2, we opted for a ceiling of 5,000 Hz to improve accuracy (Boersma & Weeninck, 2018). Further research is required to establish how best to adjust the settings by recording for formant extraction in choral singing.

Initial investigation has shown that formants in the choral signal are robust to Sound-to-Noise ratio (SNR) in line with Rathcke, Stuart-Smith, Torsney, and Harrington (2017). This meant that despite additional noise in the signal created by multiple people singing together, the formant frequencies which we usually use to evaluate vowel quality in speech were also robust enough to be used to identify vowel quality in these choral corpora.

Figures 1 and **2** are examples of Praat formant tracking over stretches of choral singing. The first example is a particularly good example from a high quality recording. F1 and F2 vary as we would expect based on the vowel qualities noted in the TextGrid. The second example is less clean. The recording quality is lower than the previous example and there are clearly some artefacts in the higher formants. However, for the most part, F1 and F2 look sensible. And generally, even where recordings are not high quality, F1 and F2 appear to be reasonably well estimated.
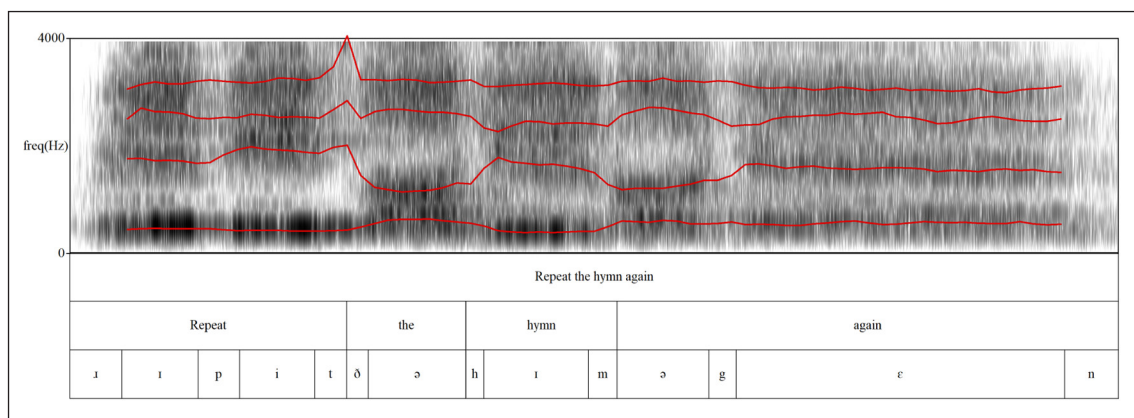

Figure 1: Example spectrogram with formant tracks of "Repeat the hymn again" from "A Great and Mighty Wonder" (King's, A Festival of Lessons and Carols, 1964, directed by David Willcocks), showing good estimation of first and second formants in Praat. Measures were taken for the underlined instances of the FLEECE, KIT, DRESS, and SCHWA vowels.
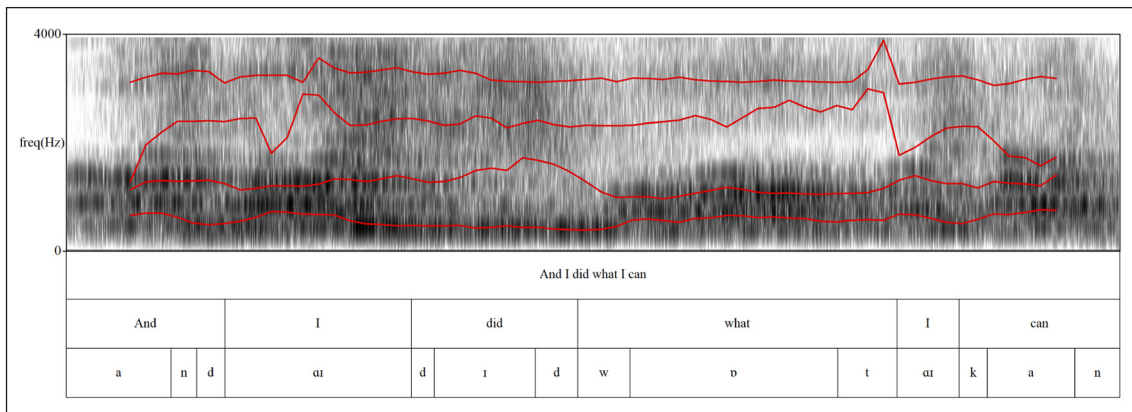
Figure 2: Example spectrogram with formant tracks of "And I did what I can" from "Remember, O Thou Man" (King's, A Festival of Lessons and Carols, 2008, directed by Stephen Cleobury), from a poorer-quality recording. Measures were taken for the underlined instances of the KIT, TRAP, and LOT vowels.

2.7 Normalisation

Tokens of less than 50 ms whose quality might be reduced were removed prior to vowel normalisation and statistical trimming following Dodsworth (2013). In addition, vowel duration was log-transformed using natural logarithm and mean centered for all models reported.

Automatically extracted F1 and F2 values were normalised using the Lobanov method (Lobanov, 1971). Lobanov (Z-score) normalisation reduces speaker-specific information in the signal, or in this case, recording-specific information that we cannot control for, e.g., type of microphone, acoustic of recording venue, sound engineer. The Lobanov method was selected because it has been found to minimise speaker-specific information while maximizing sociolinguistic variation in the signal (Rathcke et al., 2017). **Figure 3** is a plot of the vowel formant data (Lobanov normalised) after trimming (See section 2.8). The distribution looks very similar to the kind of distribution of monophthongal qualities we would expect for spoken English.

2.8 Data trimming

Following Sóskuthy and Stuart-Smith (2020), vowel tokens with F1/F2 values that fall outside the 1st and 99th percentiles for each choir were excluded. Subsequently, tokens with F1/F2 values more than 1.5 IQR (inter-quartile range) away from the lower or upper quartiles for a given vowel for each choir were also removed (Sóskuthy & Stuart-Smith, 2020). Trimming was conducted within each vowel category for F1 and F2.[4]

---

[4] We were concerned that f0 might impact the vowel formants, and we did not include f0 in the modelling reported here. However, as can be seen in Appendix D, any effect of f0 was absent in violin plots of the normalised formant data.
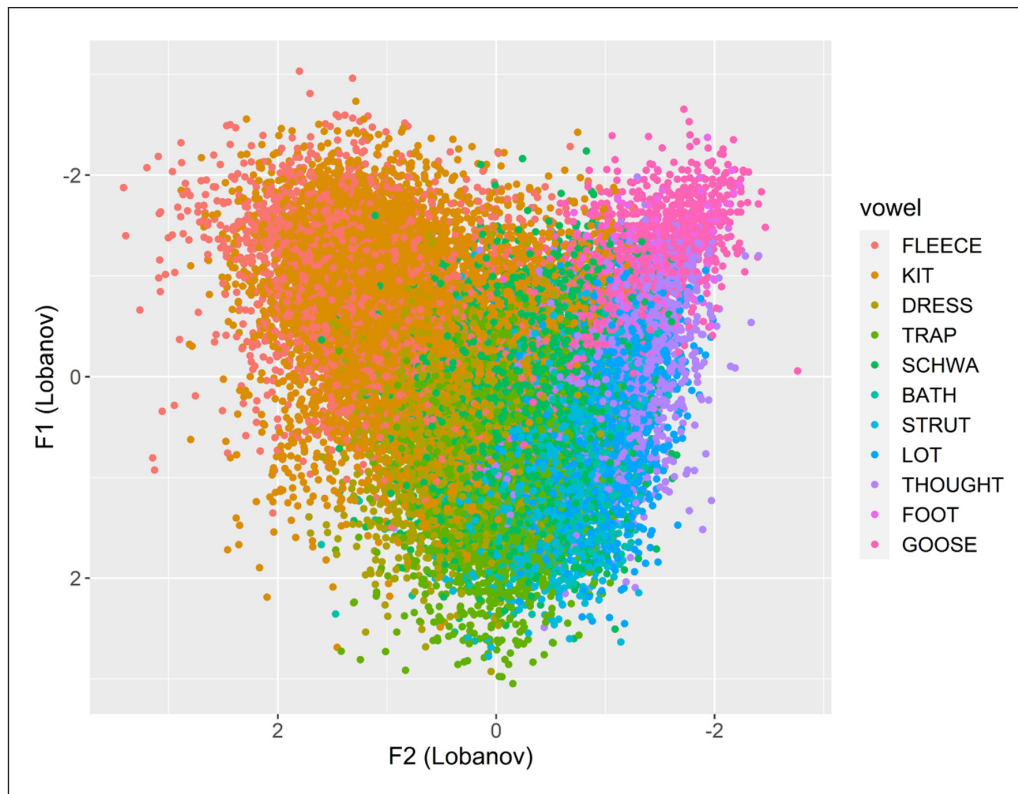
Figure 3: Acoustic vowel qualities of all vowel monophthongs from both corpora Lobanov normalised and trimmed (N = 27,432).

3. Statistical Modelling

F1 and F2 were modelled separately with Bayesian linear mixed effects models using the `brms` package (Bürkener, 2018) in R (R Core Team, 2021). We will now outline the variables used in the models which are summarised in **Table 3**.

3.1 Fixed effects:

- *Vowel*: For the front vowel analyses the levels are FLEECE, KIT, DRESS and TRAP. For the TRAP–BATH models the levels are TRAP and BATH.
- *Time/Director*: Time was grouped into three periods HSR 1925–1951, PM 1957–1975, MJS 1987–2016 for Glasgow corpus and four periods BO 1945–1957, DW 1957–1974. PL 1974–1982, and SC 1982–2019 for King's.
- *Vowel Duration*: Vowel duration was log-normalised and mean centered.
- *Genre*: Genre was coded differently for the two corpora based on the distribution in the data. For Glasgow, it was a three-level factor: Church Music, Scottish Vernacular, and Miscellaneous. King's genre was also coded as a three-level factor: Carols, Church Music, and Miscellaneous. These had to be collapsed for joint modelling purposes to a three-level

factor: Carols, Church Music, and Miscellaneous (containing Scottish Vernacular). This was due to the distribution of the number of tokens across different genres.

- *Following Segment*: Included as it is likely to influence vowel quality. As the effect of following segment in our data did not pattern in a linguistically expected way, we did not collapse the factor by place of articulation, voicing, or manner, but included a factor level for each type of segment.

- *Preceding Manner of Articulation*: In trialling with the Glasgow data, there was insufficient evidence for an effect of preceding manner of articulation, so it was removed from subsequent models.

3.2 Varying effects

There are three sources of structured variability in the data that are accounted for in the model structure:

- *Word*: is equivalent to the linguistic varying effect for "item." As this is corpus data, there are repetitions of words, and the number of tokens of each word is unbalanced. Consequently, word variances are pooled.

- *Song*: is the title of the work being sung (whether that is a song or a hymn, etc.). This varying effect functions similarly to word as a linguistic "item." It may help to think of this as a passage of text selected from a group of texts. Some "songs" are recorded multiple times over time in each corpus (e.g., "All in the April evening" in Glasgow, and "Once in Royal David's City" for King's), and some songs are recorded only once.

- *Album*: in this context is the title of a commercial release or broadcast. Album functions as a participant varying effect. There are many unknowns in the corpus data including variables such as microphone, room acoustic, recording engineer, reverb added, and medium. Album serves as a catch-all varying effect for these sources of unknown variability. Coding for all of these possible sources of systematic variability separately would be optimal but complete metadata is not available for all recordings.

These varying effects are also nested as each Word belongs to a Song, and each Song belongs to an Album. While the full nesting term (Album:Song:Word) would not harm the model due to shrinkage, it increased computational burden considerably so it was omitted.

3.3 Priors

Following results from extensive trialling with the Glasgow data, we decided to proceed with weakly informative regularising priors: normal (0, 1) for fixed effects and student_t(3, 0, 1) for varying effects as recommended by Gelman (2020). All categorical variables were sum coded for ease of interpreting model coefficients.

The following tables were produced using the `xtable` package (Dahl, Scott, Roosen, Magnusson, & Swinton, 2019), and `BayesPostEst` (Karreth, Scogin, Williams, & Beger, 2021) via `texreg` (Leifeld, 2013). Plots were produced using `brms` and `ggplot2` (Wickham, 2016). Post hoc comparisons (supplementary materials: Appendix A, Tables 1–6) were conducted using `emmeans` (Lenth, 2021).

4. Results

This paper reports results for front vowels. After statistical trimming, the resulting N were King's – FLEECE 2,010: KIT 4,675: DRESS 1,565: TRAP 2,042; Glasgow – FLEECE 902: KIT 1,941: DRESS 681: TRAP 1,003. The empirical Hz data for F1 and F2 are shown in **Tables 4** and **5**. As also found in acoustic analysis of popular singing (Gibson & Bell, 2012), we too find that the acoustic vowel quality in choral singing shows higher F1 and lower F2 values for these vowels with respect to spoken SSBE and SSE (see **Table 3** in Ferragne & Pellegrino, 2010). This results from singing technique related to enhanced resonance and projection (e.g., jaw lowering). Lobanov normalised F1 and F2 across both datasets were each analysed with Bayesian linear mixed models using `brms` in R with weakly informative priors.

Predictor	Type	Levels/Units
Time/Director	Factor	Glasgow: HSR (1925–1951), PM (1959–1975), MJS (1987–2016). King's: BO (1945–1958), DW (1959–1974), PL (1976–1982), SC (1984–2019)
Vowel	Factor	FLEECE, KIT, DRESS, TRAP, BATH
Vowel Duration	Continuous	Log ms
Following Segment	Factor	20–22 levels – See supplementary materials (Appendix C)
Genre	Factor	Carols, Church Music, Miscellaneous
Vowel:Time/Director	Interaction	
Vowel:Duration	Interaction	
Time/Director:Duration	Interaction	
Genre:Duration	Interaction	
(1|Word)	Varying effect	
(1|Song)	Varying effect	
(1|Album)	Varying effect	
(1|Song:Word)	Nesting effect	
(1|Album:Song)	Nesting effect	

Table 3: Fixed effects, interactions, and varying effects structure for modelling.

Vowel : Time/Director	N	F1 (Hz)		F2 (Hz)		Duration (ms)	
		mean	sd	mean	sd	mean	sd
FLEECE : BO 1945–1958	285	493	86	1713	277	450	320
FLEECE : DW 1959–1974	495	472	92	1687	221	600	560
FLEECE : PL 1976–1982	304	499	84	1666	229	530	410
FLEECE : SC 1984–2019	836	500	92	1751	228	740	780
KIT : BO 1945–1958	655	529	85	1572	199	360	480
KIT : DW 1959–1974	1186	491	85	1616	210	470	520
KIT : PL 1976–1982	706	519	88	1604	200	470	740
KIT : SC 1984–2019	1939	532	99	1652	212	510	530
DRESS : BO 1945–1958	229	660	84	1465	150	500	630
DRESS : DW 1959–1974	351	602	73	1501	152	620	560
DRESS : PL 1976–1982	231	637	70	1476	123	600	780
DRESS : SC 1984–2019	682	692	76	1477	136	750	570
TRAP : BO 1945–1958	277	670	94	1398	119	310	220
TRAP : DW 1959–1974	487	655	85	1391	137	410	370
TRAP : PL 1976–1982	304	671	88	1342	111	380	230
TRAP : SC 1984–2019	896	731	86	1340	105	460	450

Table 4: Raw means and standard deviations for King's front vowels by Time/Director.

Vowel : Time/Director	N	F1 (Hz)		F2 (Hz)		Duration (ms)	
		mean	sd	mean	sd	mean	sd
FLEECE : HSR 1925–1951	131	424	91	1961	93	1420	1330
FLEECE : PM 1959–1975	497	440	91	1904	116	1210	980
FLEECE : MJS 1987–2016	234	426	67	2005	147	1220	1030
KIT : HSR 1925–1951	321	446	95	1889	140	810	860
KIT : PM 1959–1975	1030	471	96	1813	128	730	580
KIT : MJS 1987–2016	414	499	104	1793	192	610	490
DRESS : HSR 1925–1951	98	629	77	1628	132	1020	920
DRESS : PM 1959–1975	406	637	81	1507	131	910	630
DRESS : MJS 1987–2016	141	677	96	1517	129	770	760
TRAP : HSR 1925–1951	144	636	100	1477	164	850	970
TRAP : PM 1959–1975	547	658	95	1445	134	690	560
TRAP : MJS 1987–2016	232	712	102	1381	135	710	650

Table 5: Raw means and standard deviations for Glasgow front vowels by Time/Director.

4.1 Model formulae

The model formulae were:

```
Vowel Formant ~ Vowel + Time + Duration + FollowingSegment + Genre
+ Vowel:Time + Vowel:Duration + Time:Duration + Genre:Duration +
(1|Album) + (1|Song) + (1|Word) + (1|Album:Song) + (1|Song:Word)
```

4.2 Model convergence criteria

For all models reported (summarised in **Tables 6** and **8**), model chains were visually inspected for convergence, Rhat was 1 for all coefficients and the minimum effective sample size for beta coefficients was greater than 400 (100 × number of chains). Posterior predictive checks for all models can be found in supplementary materials: Appendix B. We were satisfied that the models converged successfully and that the posterior summaries were amenable to interpretation.

4.3 Front vowel model summaries

The results from the modelling of F1 and F2 in the front vowels for King's and the Glasgow choirs are shown in **Table 6**. We present the results for the two formants in the following two sections.

4.4 Front vowel F1 results

Table 6 shows strong evidence of a main effect of Vowel for front vowel height, with KIT higher than the grand mean (median –0.62, CI [–0.66; –0.59]) and DRESS (median 0.50, CI [0.45; 0.54]) and TRAP (median 0.98, CI [0.93; 1.03]) lower than the grand mean. There is also evidence of a main effect of Time/Director (visualised in **Figure 4**) with Boris Ord (1945–1958, King's) (median 0.17, CI [0.01; 0.34]), Stephen Cleobury (1984–2019, King's) (median 0.15, CI [0.05; 0.25]), and Marilyn J. Smith (1987–2016, Glasgow) (median 0.21, CI [0.06; 0.36]) patterning together lower than the grand mean. David Willcocks (1959–1974, King's) is more raised than the grand mean (median –0.31, CI [–0.44; –0.18]).

There is evidence of an interaction of Vowel by Time/Director. This interaction is visualised in **Figure 4**. Post hoc comparisons reported here can be found in supplementary materials: Appendix A Table 1. As can be seen in **Figure 4**, Hugh S. Roberton (1925–1951, Glasgow) and David Willcocks (1959–1974, King's) largely pattern together in front vowel height and post hoc comparisons show that there are no differences between them for FLEECE, KIT and TRAP. There is, however, a difference for DRESS with recordings made under Willcocks exhibiting a much more raised realisation than under Roberton (median –0.39, CI [–0.68; –0.13]) closer to [e] than [ɛ]. Peter Mooney (1959–1975, Glasgow) and Philip Ledger (1976–1982, King's) both follow Roberton and Willcocks respectively and show a modest lowering for all front vowels from the peak of acoustic front vowel height under both choirs' previous leaders.

	F1 Est	95% CI		F2 Est	95% CI	
Intercept	–0.02	–0.09	0.05	0.58	0.52	0.64
vowelKIT	**–0.62**	**–0.66**	**–0.59**	**0.33**	**0.30**	**0.36**
vowelDRESS	**0.50**	**0.45**	**0.54**	**–0.29**	**–0.33**	**–0.25**
vowelTRAP	**0.98**	**0.93**	**1.03**	**–0.70**	**–0.75**	**–0.66**
directorPM1959–1975	0.00	–0.12	0.13	0.04	–0.08	0.15
directorMJS1987–2016	**0.21**	**0.06**	**0.36**	0.08	–0.05	0.21
directorBO1945–1958	**0.17**	**0.01**	**0.34**	**–0.18**	**–0.33**	**–0.02**
directorDW1959–1974	**–0.31**	**–0.44**	**–0.18**	–0.03	–0.15	0.09
directorPL1976–1982	–0.09	–0.26	0.07	–0.14	–0.30	0.02
directorSC1984–2019	**0.15**	**0.05**	**0.25**	–0.01	–0.10	0.09
Vowel Duration (log)	–0.00	–0.02	0.01	**–0.02**	**–0.04**	**–0.01**
vowelKIT : directorPM1959–1975	–0.00	–0.04	0.04	**0.10**	**0.06**	**0.13**
vowelDRESS : directorPM1959–1975	0.03	–0.02	0.09	**–0.12**	**–0.16**	**–0.07**
vowelTRAP : directorPM1959–1975	**–0.12**	**–0.17**	**–0.07**	0.02	–0.02	0.06
vowelKIT : directorMJS1987–2016	–0.05	–0.11	0.01	**0.09**	**0.04**	**0.13**
vowelDRESS : directorMJS1987–2016	**0.17**	**0.09**	**0.25**	**–0.18**	**–0.24**	**–0.11**
vowelTRAP : directorMJS1987–2016	0.07	–0.00	0.14	**–0.12**	**–0.18**	**–0.06**
vowelKIT : directorBO1945–1958	0.04	–0.01	0.09	**–0.21**	**–0.26**	**–0.17**
vowelDRESS : directorBO1945–1958	–0.06	–0.13	0.00	**0.06**	**0.01**	**0.12**
vowelTRAP : directorBO1945–1958	0.01	–0.05	0.08	**0.16**	**0.10**	**0.22**
vowelKIT : directorDW1959–1974	**0.04**	**0.00**	**0.08**	**–0.06**	**–0.10**	**–0.03**
vowelDRESS : directorDW1959–1974	**–0.19**	**–0.24**	**–0.13**	**0.10**	**0.06**	**0.15**
vowelTRAP : directorDW1959–1974	0.02	–0.04	0.07	**0.11**	**0.06**	**0.15**
vowelKIT : directorPL1976–1982	0.03	–0.02	0.07	–0.02	–0.06	0.02
vowelDRESS : directorPL1976–1982	**–0.09**	**–0.16**	**–0.03**	**0.13**	**0.08**	**0.19**
vowelTRAP : directorPL1976–1982	–0.01	–0.07	0.05	–0.02	–0.07	0.03
vowelKIT : directorSC1984–2019	**–0.09**	**–0.12**	**–0.06**	**0.03**	**0.00**	**0.06**
vowelDRESS : directorSC1984–2019	**0.07**	**0.03**	**0.12**	**0.05**	**0.01**	**0.09**
vowelTRAP : directorSC1984–2019	**0.11**	**0.06**	**0.15**	**–0.14**	**–0.18**	**–0.11**
vowelKIT : Vowel Duration (log)	**–0.09**	**–0.11**	**–0.07**	**0.09**	**0.07**	**0.10**
vowelDRESS : Vowel Duration (log)	**0.07**	**0.04**	**0.10**	**–0.08**	**–0.10**	**–0.05**
vowelTRAP : Vowel Duration (log)	**0.12**	**0.09**	**0.14**	**–0.08**	**–0.10**	**–0.06**

Table 6: Combined front vowel F1 and F2 model posterior summaries. Estimates for `Genre`, `Genre:VowelDuration`, and `Time/Director:VowelDuration` can be inspected from the models available on the OSF https://osf.io/3vjxm/. **Bold type** indicates 0 outside 95% credible interval.
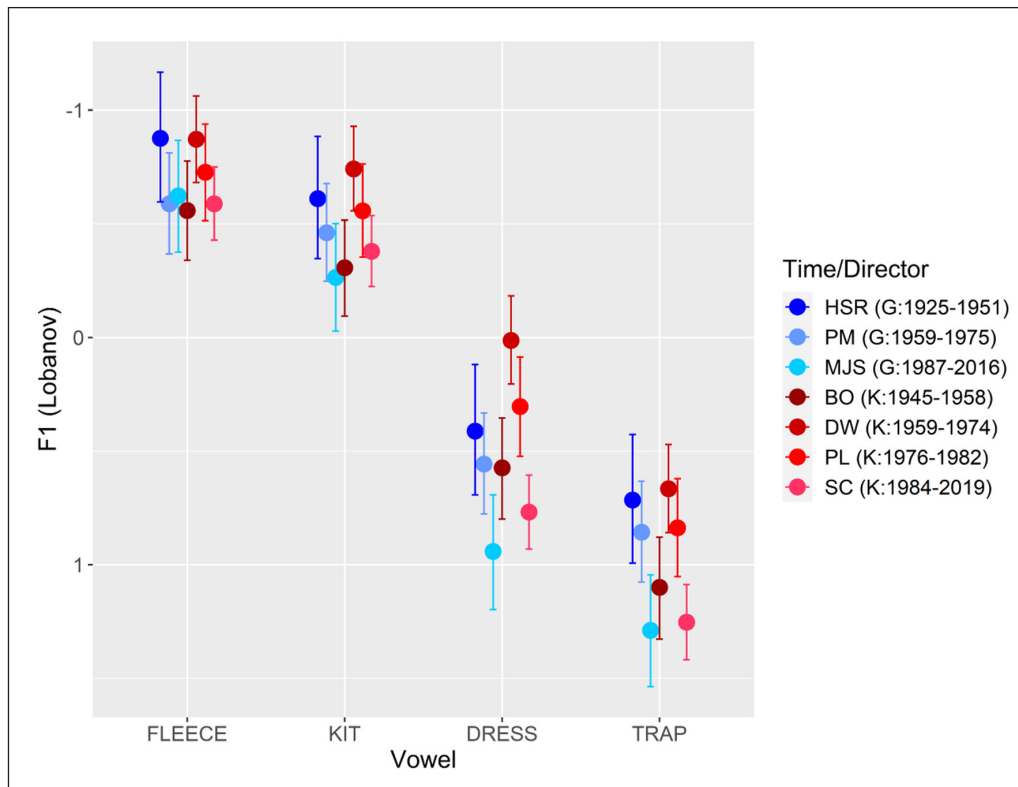
Figure 4: Combined front vowel F1 model estimates for Vowel by Time/Director interaction (bars show 95% credible intervals; G = Glasgow corpus, K = King's corpus; N = 14,404).

Boris Ord, Stephen Cleobury and Marilyn J. Smith pattern together in front vowel height for FLEECE, KIT, and TRAP, and post hoc tests reveal no differences between them. There is a difference between Boris Ord and Marilyn J. Smith for DRESS with Ord being more raised (median –0.37, CI [–0.63; –0.09]). Broadly, in terms of front vowel height, the later Time/ Director pairs (Stephen Cleobury, and Marilyn J. Smith) for both King's and Glasgow produce a similar acoustic vowel quality to that produced under Boris Ord.

Returning to the posterior summary in **Table 6**, there is no evidence of a main effect of Vowel Duration or Genre for the combined front vowel height model. As expected, there is evidence of an interaction of Vowel by Vowel Duration, such that as Vowel Duration increases, KIT raises (median –0.09, CI [–0.11; –0.07]) while DRESS (median 0.07, CI [0.04; 0.10]) and TRAP (median 0.12, CI [0.09; 0.14]) lower.

Thus, the key finding in terms of vowel height is that the front vowels have lowered in both the Glasgow and King's data. There is also some evidence that the choirs have become more similar over time in acoustic vowel quality.

4.5 Front vowel F2 results

The posterior summary for the combined front vowel F2 model can be found in **Table 6**. As expected, there is a strong main effect of Vowel, with KIT being fronter than the grand mean (median 0.33, CI [0.30; 0.36]) and DRESS (median –0.29, CI [–0.33; –0.25]) and TRAP (median –0.70, CI [–0.75; –0.66]) backer than the grand mean. There is a limited main effect of Time/Director with Boris Ord being more retracted than the grand mean for Time/Director (median –0.18, CI [–0.33; –0.02]).

There is also an interaction of Vowel by Time/Director. The interaction is visualised in **Figure 5**. Post hoc comparisons (supplementary materials: Appendix A, Table 2) show that FLEECE and KIT have fronted under Stephen Cleobury and there are no differences between Stephen Cleobury and the three Glasgow time periods. KIT has likewise fronted for Stephen Cleobury and there are no differences with Peter Mooney and Marilyn J. Smith apart from being different to Hugh S. Roberton (median –0.29, CI [–0.51; –0.06]). For DRESS, Stephen Cleobury is more retracted than Marilyn J. Smith (median 0.21, CI [0.02; 0.41]), but not different to Peter Mooney or Hugh S. Roberton. For TRAP, Stephen Cleobury is more retracted than Hugh S. Roberton (median –0.37, CI [–0.59; –0.12]) and Peter Mooney (median –0.18, CI [–0.35; –0.01]), but there is no difference to Marilyn J. Smith.
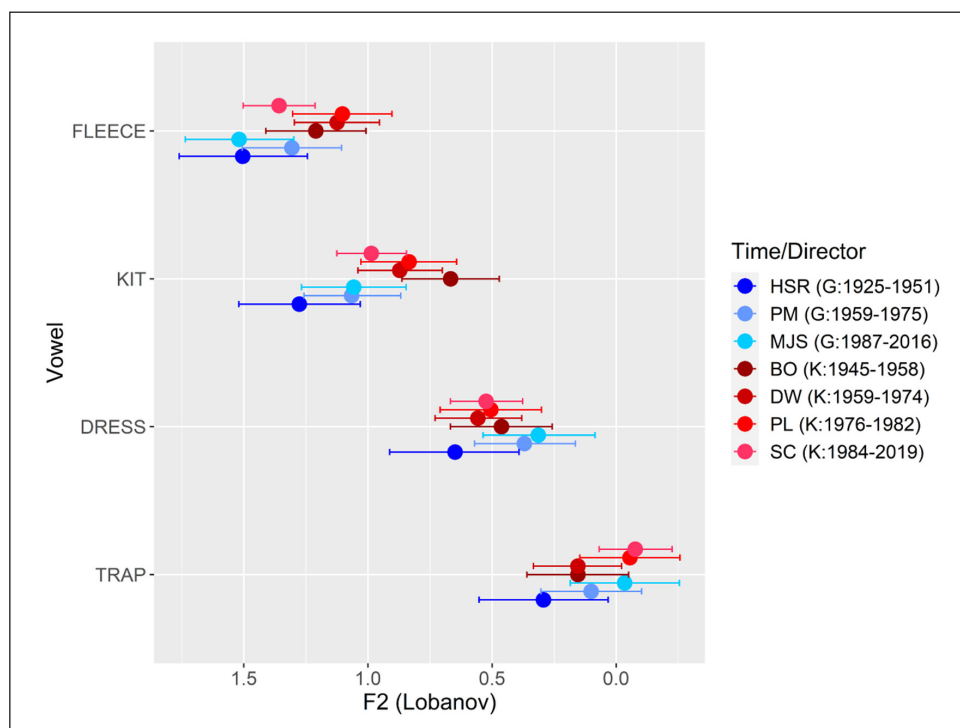

Figure 5: Combined front vowel F2 model estimates for Vowel by Time/Director interaction (bars show 95% credible intervals; G = Glasgow corpus, K = King's corpus; N = 14,404).

Returning to the combined front vowel F2 posterior summary (**Table 6**), in contrast to the F1 model, there is evidence supporting main effects of both Vowel Duration and Genre. The greater Vowel Duration, the more front vowels retract overall (median −0.02, CI [−0.04; −0.01]). The Genre "Church Music" is more retracted than the grand mean for Genre (median −0.08, CI [−0.15; −0.02]).

As predicted, there is an interaction of Vowel by Vowel Duration, such that as Vowel Duration increases, KIT fronts (median 0.09, CI [0.07; 0.10]) while DRESS (median −0.08, CI [−0.10; −0.05]) and TRAP (median −0.08, CI [−0.10; −0.06]) retract. There was also an interaction of Genre by Vowel Duration. As Vowel Duration increases, vowels in the Genre "Church Music" retract more (median −0.03, CI [−0.05; −0.02]).

In short, the main finding for F2 is that the realisations of FLEECE, KIT, and TRAP in Glasgow and King's have become more similar over time. DRESS behaves somewhat differently with the Glasgow realisation retracting over time to a more central position than at King's. In contrast, DRESS has stayed fairly stable at King's over time.

Vowel : Time/Director	N	F1 (Hz)		F2 (Hz)		Duration (ms)	
		mean	**sd**	**mean**	**sd**	**mean**	**sd**
TRAP : HSR (1925–1957)	160	627	102	1488	172	770	940
TRAP : PM (1959–1975)	637	650	98	1450	136	660	580
TRAP : MJS (1987–2016)	262	708	109	1383	151	640	650
TRAP : BO (1945–1958)	277	689	94	1398	119	310	220
TRAP : DW (1959–1974)	487	654	85	1390	137	410	370
TRAP : PL (1976–1982)	304	671	88	1342	111	380	230
TRAP : SC (1984–2019)	896	730	86	1339	105	460	450
BATH : HSR (1925–1957)	63	672	77	1336	251	950	850
BATH : PM (1959–1975)	166	639	90	1258	136	1060	1000
BATH : MJS (1987–2016)	78	661	94	1208	101	760	530
BATH : BO (1945–1958)	116	703	65	1176	75	1090	2140
BATH : DW (1959–1974)	142	652	85	1139	105	990	470
BATH : PL (1976–1982)	88	670	73	1178	97	980	1580
BATH : SC (1984–2019)	254	709	76	1180	96	1270	2110

Table 7: Combined TRAP–BATH raw N, means and standard deviations.

	F1 Est	95% CI		F2 Est	95% CI	
Intercept	**0.73**	**0.61**	**0.86**	**–0.69**	**–0.76**	**–0.62**
vowelTRAP	**0.24**	**0.14**	**0.35**	**0.49**	**0.43**	**0.55**
directorPM1959–1975	–0.09	–0.28	0.10	0.09	–0.02	0.19
directorMJS1987–2016	0.21	–0.03	0.45	0.01	–0.11	0.14
directorBO1945–1958	0.23	–0.01	0.46	–0.00	–0.13	0.13
directorDW1959–1974	**–0.29**	**–0.49**	**–0.10**	**–0.12**	**–0.23**	**–0.02**
directorPL1976–1982	–0.23	–0.48	0.02	–0.07	–0.20	0.07
directorSC1984–2019	**0.18**	**0.02**	**0.34**	–0.04	–0.12	0.05
Vowel Duration (log)	–0.01	–0.06	0.04	**–0.09**	**–0.12**	**–0.05**
vowelTRAP : directorPM1959–1975	–0.01	–0.13	0.10	–0.06	–0.13	0.01
vowelTRAP : directorMJS1987–2016	0.08	–0.08	0.24	**–0.13**	**–0.23**	**–0.04**
vowelTRAP : directorBO1945–1958	–0.04	–0.20	0.11	**0.14**	**0.05**	**0.23**
vowelTRAP : directorDW1959–1974	–0.04	–0.17	0.09	**0.27**	**0.20**	**0.34**
vowelTRAP : directorPL1976–1982	0.05	–0.11	0.20	**–0.11**	**–0.20**	**–0.02**
vowelTRAP : directorSC1984–2019	**0.13**	**0.02**	**0.23**	**–0.09**	**–0.15**	**–0.03**
vowelTRAP : Vowel Duration (log)	**0.12**	**0.06**	**0.17**	–0.02	–0.05	0.02

Table 8: Combined TRAP–BATH F1 and F2 model posterior summaries. Estimates for `Genre, Genre:Vowel Duration (log)` and `Time/Director:Vowel Duration` can be inspected from the models available on the OSF https://osf.io/3vjxm/. **Bold type** indicates 0 outside 95% credible interval.

4.6 Combined TRAP–BATH model

Recall from Section 1.6 that Scottish English shows a single vowel for the lexical sets TRAP and BATH, whereas SSBE shows the TRAP–BATH split. If the Glasgow choir singing is based on an SSE phonology then we would expect no TRAP–BATH split, in contrast to King's where we would expect to find TRAP–BATH split. The main finding shown in **Tables 7** and **8** is that there is a TRAP–BATH split present in both the King's and Glasgow data in F2.

Estimates from the combined TRAP–BATH height model can be found in **Table 8**. There is evidence of a main effect of Vowel with TRAP lower in height than the grand mean (median 0.24, CI [0.14; 0.35]). There is also evidence of a main effect of Time/Director with David Willcocks (1959–1974, King's) being higher than the grand mean (median –0.29, CI [–0.49; –0.10]) and Stephen Cleobury (1984–2019, King's) being lower than the grand mean (median 0.13, CI [0.02; 0.23]).

There is an interaction of Vowel by Time/Director driven by Stephen Cleobury's TRAP vowel realisation being considerably lower than the grand mean (median 0.13, CI [0.02; 0.23]). This interaction is visualised in **Figure 6**. Post hoc comparisons can be found in supplementary materials (Appendix A) – by Vowel (Table 3) – and by Time/Director (Table 4). Stephen Cleobury, Marilyn J. Smith and Boris Ord do not differ in height for TRAP or BATH. There is a difference between TRAP and BATH height for Marilyn J. Smith (median 0.22, CI [0.02; 0.42]) and Stephen Cleobury (median 0.27, CI [0.13; 0.40]) but not for other Time/Director pairs. Therefore, TRAP–BATH has become statistically distinguishable in height over time in both Glasgow and King's corpora. However, it is not clear how this relates to perception and whether these differences would be noticeable to naive listeners.

Figure 6: Combined TRAP–BATH F1 model estimates for Vowel by Time/Director interaction (bars show 95% credible intervals; G = Glasgow corpus, K = King's corpus; N = 3,928).

For the TRAP–BATH F1 model, there is no evidence of a main effect of Vowel Duration or Genre. There is evidence for an interaction of Vowel by Vowel Duration with TRAP lowering as duration increases (median 0.12, CI [0.06; 0.17]).

We will now turn our attention to the F2 dimension. The main difference in acoustic quality in TRAP–BATH in SSBE is a difference in F2, with BATH being more retracted /ɑː/ and TRAP

being fronter. If there is a statistical difference to be found between TRAP–BATH for the Scottish corpus we would expect to find it in the F2 dimension. If the Scottish choral accent is based on a Standard Scottish English phonology we would expect to find next to no difference between TRAP–BATH in F2, as they should have a singular phoneme. If the phonology is based on Southern Standard British English, then we would expect to find a distinction in F2 between TRAP–BATH in both the Glasgow and King's corpora.

The combined TRAP–BATH F2 model summary can be found in **Table 8**. There is strong evidence of a main effect of Vowel with TRAP substantially fronter than the grand mean (median 0.49, CI [0.43; 0.55]), as expected. There is also a main effect of Time/Director with David Willcocks (1959–1974, King's) more retracted than the grand mean (median –0.12, CI [–0.23; –0.02]) overall. There is evidence of an interaction of Vowel by Time/Director visualised in **Figure 7**. Post hoc comparisons can be found in supplementary materials (Appendix A) – by Vowel (Table 5) – by Time/Director (Table 6).
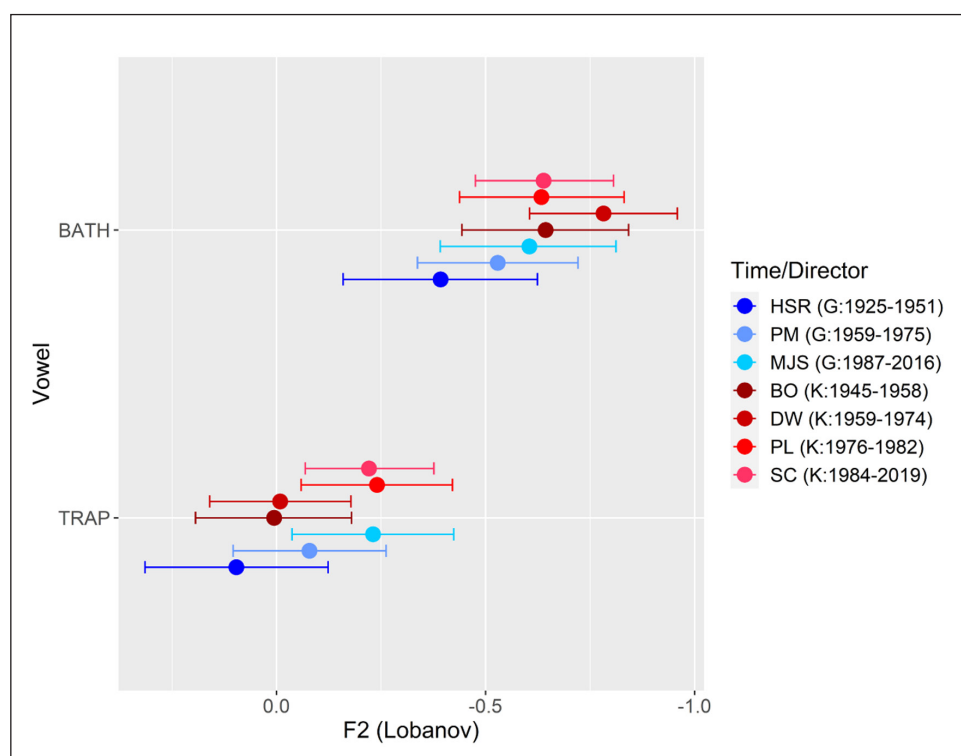

Figure 7: Combined TRAP–BATH F2 model estimates for Vowel by Time/Director interaction (bars show 95% credible intervals; G = Glasgow corpus, K = King's corpus; N = 3,928).

This analysis provides evidence of TRAP retracting over time in both Glasgow and King's datasets, with no differences shown between the late time periods in each corpus Marilyn J. Smith – Stephen Cleobury (median 0.0088, CI [–0.13; 0.15]). The Glasgow BATH vowel appears

to have retracted over time and now rests at similar backness to King's, with no difference between the later two time periods Marilyn J. Smith – Stephen Cleobury (median –0.03, CI [–0.20; 0.13]). There is a statistical difference between TRAP and BATH in the F2 domain for all Time/Director pairs, as visualised in **Figure 7**. However, the degree of F2 difference between TRAP and BATH may be decreasing over time in both Glasgow and King's, as both the estimates and width of credible intervals decrease over time.

5. Discussion

Phonetic studies of singing to date have tended to focus on the acoustics of the voice in various types of professional solo singing. Sociolinguistic studies of singing have mostly been restricted to popular singing styles such as pop, rock, punk, and heavy metal. In this study, we investigated language variation and change in recordings of choral singing. We investigated the inventory and realisation of front vowels produced by British choirs (1925–2019) in over 25 hours of unaccompanied choral singing in English, comprising 74 albums, 226 song types, and over 1,300 different word types. The choirs investigated were the Glasgow Orpheus and Phoenix choirs (FLEECE 902: KIT 1,941: DRESS 681: TRAP 1,003), and the Choir of King's College, Cambridge (FLEECE 2,010: KIT 4,675: DRESS 1,565: TRAP 2,042). Both datasets were analysed together using Bayesian linear mixed models with `brms` in R and weakly informative priors. The analyses provide evidence of change over time, and specifically a main effect of acoustic vowel lowering over time in the front vowels of both corpora.

The TRAP–BATH analysis also showed that there was a difference between TRAP and BATH in either F1 or F2 for all time periods in both corpora – which is not what we would expect if the vowel phonology of both corpora was based on the regional variety of the singers. Thus this study provides the first evidence for variation and change in choral singing, and supports a shared British choral accent, at least for vowel quality.

The following discussion is structured around the research questions in 1.6. Firstly, we will discuss how there appears to be a common front vowel system in both structure and realisation. Secondly, we will discuss how the front vowel realisations have systematically changed over time in both corpora. Following this, we relate our findings to the musicological literature.

5.1 British classical choirs show a common front vowel system

This section addresses research questions 1a and 1b (1.6). In this study, we investigated the front vowels FLEECE, KIT, DRESS and TRAP, and evidence for the TRAP–BATH split in the Glaswegian choirs. **Figure 8** gives a synchronic snapshot by plotting F1 and F2 for each vowel separated by corpus, not separated by time.
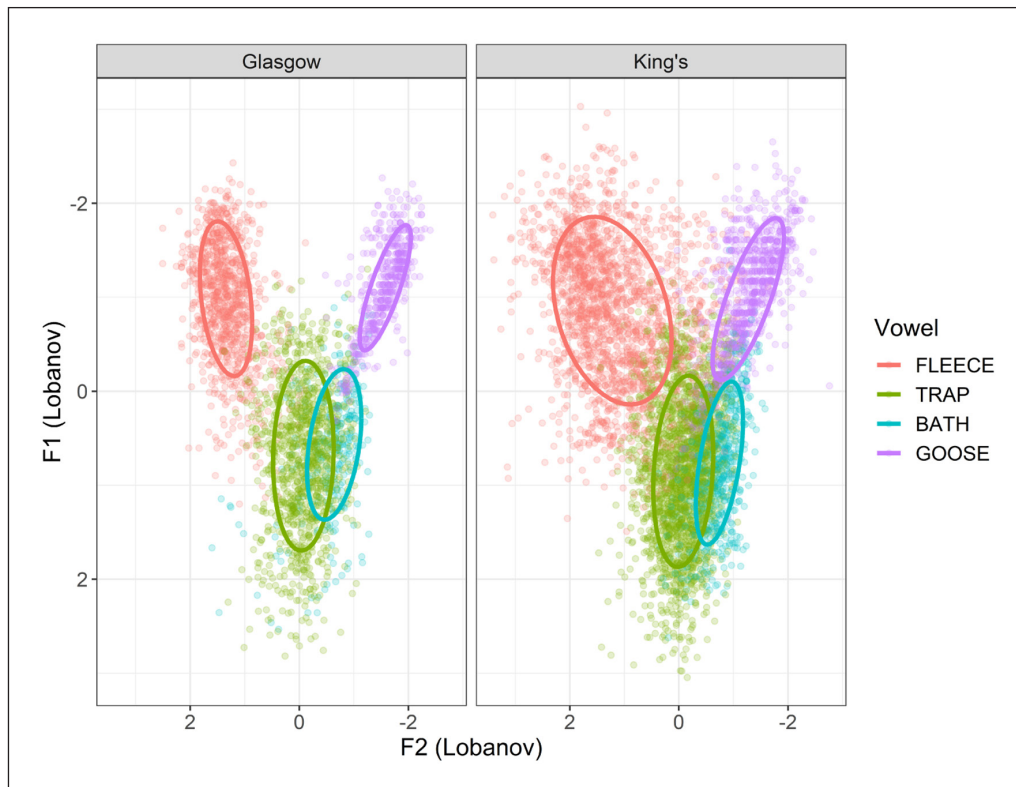
Figure 8: Vowel plots of formant measures (Lobanov) by choir corpus. Ellipses show 1 Standard Deviation from the mean; /uː/ is included for orientation.

As can be seen in **Figure 8**, the Choir of King's College, Cambridge, and the Glasgow Orpheus and Phoenix choirs have front vowel systems that look very similar in both in phonology and realisation. We note, however, that the F2 for King's is more variable than for the Glasgow recordings. We suspect that this is partly due to the different sample sizes. But, **Tables 4** and **5** show that the durations of the Glasgow vowels are substantially longer than those of King's – the King's vowel qualities may be more centralised as a result, leading to the wider distribution.

Overall, there is similarity in the way the vowel ellipses are positioned spatially in reference to each other within and across corpora. This shared system appears to be based on a Standard Southern British English (SSBE) phonology as suggested by Potter (1998), Sagrans (2016), and Day (2018) about the choral phonology of King's College. For front vowels at least, there is evidence of a shared non-regional British choral accent. The finding of separation in acoustic qualities for TRAP and BATH lexical sets in the Glasgow choirs, distinct in F2, is unexpected. This suggests that the vowel phonology of choral singing in Glasgow is at least partly based on a non-regional standard accent linked to SSBE rather than local varieties of Scottish Standard English or Glaswegian English where there is a single vowel phoneme, and hence a low vowel continuum from [a – ɑ] (Abercrombie, 1979; P. Johnston, 1997; Stuart-Smith, 2003). This evidence for the

TRAP–BATH split in the Glasgow choirs is surprising, since, as Wells comments, "RP does not enjoy the same tacit status in Scotland as it does in England or Wales" (Wells, 1982b, p. 393).

5.2 British classical choral front vowels have changed over time

In this section, we address research questions 2a and 2b (1.6). Diachronic phonetic studies of speech have showed that SSBE has changed over time and this is particularly salient for the TRAP vowel in the twentieth century with, for example, the phonetic realisation of *cat* /kat/, changing from [kæt] to [kat] (e.g., Wells, 1982a; Harrington et al., 2000; Fabricius, 2007); see also Section 5.2.1. In this study, we find evidence for a main effect of Time/Director for front vowel F1 reflecting lowering over time, as predicted, but in both choir corpora, irrespective of spoken accent. This finding is therefore also consistent with the notion that the accent of British classical choral singing is based on a Southern Standard British English.

Recordings produced under the choir directors Hugh S. Roberton (1925–1951, Glasgow) and David Willcocks (1957–1974, King's) do not statistically differ in vowel height for FLEECE, KIT, and TRAP, despite not overlapping in time. The only difference is for DRESS, for which Willcocks produces a strikingly more raised realisation, akin to [e]. This finding lends further credibility to the directors Roberton and Willcocks having an RP speech target. It also supports the connection Day (2018) draws between the sound produced by the Choir of King's College, Cambridge under the director David Willcocks and conservative-RP with the particularly raised realisation of DRESS [eɪ] (see U-RP in Wells 1982a, pp. 280–282).

Both Stephen Cleobury and Marilyn J. Smith produce a similar front vowel height overall to Boris Ord. This suggests that the predicted lowering was already complete at King's *before* David Willcocks took over as director. It is unknown whether King's had a more raised front vowel height earlier in the century as there are very few recordings before 1945.

Our analysis has provided evidence of diachronic variation and change in recordings of choral singing. There are a number of factors that might influence variation and change in choral accents which warrant further investigation, including: the singers' own spoken accents; the accent of their directors; and/or variation in the director's artistic vision/target "accent." For example, do singers imitate their director's realisations, or do they do what they are told? Recent research has investigated what singers want from their conductors (Cronie, 2021) and found that the artistic vision of the director was one of a set of expectations of the choir members. How might artistic vision relate to linguistic variation and change?

5.2.1 One factor in determining a choir's accent: The choir director

What can we infer about the influence of a choir director on a choir's sound? When David Willcocks was director of King's (1957–1974), he reintroduced the raised [kæt] variant of the TRAP vowel, which is consistent with early twentieth-century conservative-RP pronunciation.

This is similar to the front vowel height produced by the Glasgow Orpheus Choir under Roberton (in recordings from 1945–1951), which may have been modelled on RP and/or perhaps the contemporary prestigious Kelvinside/Morningside accents of Glasgow and Edinburgh, which were also known for extremely raised realisations of DRESS (P. A. Johnston, 1985). The connection between the King's style developed under Willcocks and conservative-RP is quite convincing, as Day writes:

> Inevitably, Willcocks cultivated certain sounds which reflected his own style of spoken English, perhaps more the received pronunciation of English he heard as a chorister at Westminster Abbey in the 1930s than that of the 1960s. So *alleluia* became "e-lleluia." "I know thett my Redeemer liveth, ent thett he shell stent…" (Day, 2018, p. 261).

Phonetically, just as we've seen, /alɛluːjə/ became [ælæluːjə].

A possible scenario is that front vowel lowering took place in the King's choral accent alongside the documented shift in RP and likely the choir members own accents over time. However, under Willcocks's direction, the choir reverted to the conservative-RP front vowel realisations he himself had experienced when he was a chorister at Westminster Abbey in the 1930s.

5.2.2 A British classical choral "reference style"

Finding the TRAP–BATH split in recordings of choral singing from Glasgow perhaps reflects the findings of Caillol and Ferragne (2019) relating to the FOOT–STRUT split produced by the singers of British heavy metal bands. FOOT and STRUT were both realised as [ʊ] in the spoken interviews with Def Leppard (from Sheffield). However, FOOT and STRUT were realised as [ʊ] and [ʌ] respectively in the recordings of their singing. In contrast, Iron Maiden (from London) produced the split in both interviews and sung recordings. Caillol and Ferragne (2019) suggest that this is evidence of Def Leppard adapting to a USA model of singing pronunciation. Specifically, that producing the FOOT–STRUT contrast was deemed more stylistically appropriate for the performance of heavy metal because it is more consistent with an "American" accent. This brings us back to the notion of the reference style (Morrissey, 2008). British heavy metal bands exist within the domain of the popular reference style which is largely considered to be based on "American" norms. Similarly, we suggest that British choirs exist within the domain of a classical reference style which is modelled on SSBE norms. This argument assumes Morrissey (2008), Beal (2009), and Gibson (2019)'s notions of stylistic appropriateness for a particular musical genre existing within a larger reference style. Thus, in this case, the TRAP–BATH split found in recordings of choirs in a dialect area where there is no TRAP–BATH split present in speech suggests that producing the TRAP–BATH split is stylistically appropriate for classical choral singing. British English may have become institutionalised with respect to Western classical choral singing (Wilson, 2014), and here we provide evidence from recordings of British choral singing that the choral accent is likely based on SSBE specifically.

We have established that British choirs from different dialect areas show a shared phonology, insofar as their front vowel system is concerned, and that overall change over time supports a connection between choral singing and a non-regional accent. In terms of realisation, if Glasgow choral singing is based on spoken Standard Scottish English vowel quality, then we would expect TRAP /a/ to be more central than in SSBE, and for DRESS /ɛ/ to be more raised (Wells, 1982b). What we find is that TRAP has lowered and retracted for both King's and the Glasgow choirs, following a pattern of lowering and retraction in RP/SSBE over the twentieth century (e.g., Harrington et al., 2000; Hawkins & Midgley, 2005; Fabricius, 2007; Bjelaković, 2017). King's producing a more raised DRESS quality likely reflects the conservative RP target suggested by Day (2018). Overall, however, TRAP and DRESS have converged in F1 over time in the choir corpora. FLEECE, KIT, and TRAP have converged in F2 over time in the choir corpora.

Our findings support a convergence of acoustic vowel quality for front vowels produced by the Choir of King's College, Cambridge and the Glasgow Orpheus and Phoenix choirs. Considering also the robust distinction between TRAP and BATH reported in both regional choirs, and the common front vowel lowering, there is evidence for a common standard, or reference style, for classical choral singing, based on a non-regional British spoken accent.

5.3 Reflections and directions for future research

As the first foray into the acoustic analysis of classical choir singing, there are a number of questions that remain unanswered and which require investigation in future study. In terms of technical analysis, what are best practices for formant extraction for choral singing? Whose voices are being heard in this analysis? That is, are the extracted formant data a flat average of all of the values produced by each singer, or is it biased towards a certain voice type and/or gender?

This study considered only the front vowel system. Would other phonological variables also support a common British choral accent, or are some features shared (front vowels) while others show regional dialect differences? Marshall (2023a) shows that different variables can indeed show their own trajectories, influenced by different factors (e.g., dialect, director), but overall he finds a focusing towards a non-regional British choral accent. Specifically, a major phonological difference between Southern Standard British English and Scottish Englishes is that the latter is rhotic where SSBE is not. What happens to the realisation of underlying rhotic phonology in singing? Marshall (2023b) shows that rhoticity is much more commonly realised in the Glasgow corpus than for King's, suggesting that the sung consonant phonology can be impacted by regional dialect.

Another possible direction for future work concerning change over time in choral sound would be to investigate recordings of German choirs which have a comparable history of recording.

There is some indication here that classical choral singing practice in the UK is tethered to the non-regional accent: SSBE. It is possible there is a similar situation in Germany with *Hochdeutsch* (high German) which is a comparable non-regional accent. Referring to solo singing, A. Johnston writes: "In lyric diction, one strives to sing in the 'high form' of each language. For example, in English there is Received Pronunciation (RP) for Great Britain and General American English (GA) for the United States and Canada; in French one uses Parisian French; in German one uses *Hochdeutsch*" (A. Johnston, 2016, p. 147). If there are vowel contrasts present in Hochdeutsch that are absent in a certain regional variety, it would be interesting to see if this contrast is found in the singing of speakers from regions which do not have the contrast in speech, as we have found in this research.

6. Conclusion

This study provides the first empirical quantitative evidence supporting the connection between classical choral singing and Southern Standard British English. Our results show a shift in vowel height in British choral singing, with KIT, DRESS, and TRAP lowering over time, mirroring diachronic phonetic studies of RP (e.g., Harrington et al., 2000; Bjelaković, 2017). There is evidence for the TRAP–BATH split present in recordings of choral singing from both Cambridge and the Glasgow choirs, where spoken Scottish English maintains a single vowel phoneme; there is also evidence for a convergence of acoustic vowel quality across the two choral datasets. This suggests that SSBE features are stylistically appropriate for the performance of British classical choral singing and are perhaps features of a wider classical choral style (Wilson, 2014; Wilson, 2017).

Front vowel lowering was already well–advanced in the King's choir recordings of the 1950s. However, when David Willcocks became director (1959–1974), he raised the vowel quality, "[cultivating] certain sounds which reflected his own style of spoken English, perhaps more the received pronunciation of English he heard as a chorister at Westminster Abbey in the 1930s than that of the 1960s" (Day, 2018, p. 261). The front vowel height heard in more recent recordings of King's then returned to its original trajectory of lowering, first observed in the 1950s under Boris Ord. This provides evidence of the impact a particular director and their vision can have on variation and change in a choir's sound. Future research is needed to investigate the extent to which other phonological variables are amenable to choral direction and/or regional variation in classical British choir singing, and more generally for choral accents.

Additional file

The additional file for this article can be found as follows:

- **Supplementary materials.** A Choral vowel formant model post hoc comparisons. DOI: https://doi.org/10.16995/labphon.10125.s1

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors have made substantial contributions to the design and interpretation of this work. The first author annotated and extracted the data. Data analysis was performed by the first and second author. All authors contributed significantly to the interpretation of the results. The manuscript was drafted by the first and second author and revised by the other authors.

References

Abercrombie, D. (1979). The accents of Standard English in Scotland. In A. J. Aitken & T. MacArthur (Eds.), *Languages of Scotland* (p. 68–84). Chambers of Edinburgh.

Adams, D. (2008). *A handbook of diction for singers: Italian, German, French* (Second ed.). Oxford University Press.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). Celex2 (Computer software manual No. LDC96L14.). Philadelphia: Linguistic Data Consortium. DOI: https://doi.org/10.35111/gs6s-gm48

Beal, J. C. (2009). "You're not from New York City, you're from Rotherham" Dialect and identity in British indie music. *Journal of English Linguistics, 37*, 233–240. DOI: https://doi.org/10.1177/0075424209340014

Bjelaković, A. (2017). The vowels of contemporary RP: Vowel formant measurements for BBC newsreaders. *English Language and Linguistics, 21*(3), 501–532. DOI: https://doi.org/10.1017/S1360674316000253

Boersma, P., & Weeninck, D. (2018). Praat: doing phonetics by computer [computer program]. version 6.0.43 [Computer software manual]. Retrieved from http://www.praat.org/

Bürkener, P. C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal, 10*, 395–411. DOI: https://doi.org/10.32614/RJ-2018-017

Burns, A., & Kydd, C. (2013). *Choir leader's training manual.* Scotland Sings – Hands up for Trad.

Caillol, C., & Ferragne, E. (2019). The sociophonetics of British heavy metal music: T voicing and the FOOT-STRUT split. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings*

of the 19th International Congress of Phonetic Sciences, Melbourne, Australia. Australasian Speech Science and Technology Association, Inc. Retrieved from https://assta.org/proceedings/ICPhS2019Microsite/

Cronie, K. (2021). *"Voice-centred choral conductorship": An exploration of singers' expectations of uk-based choral leaders* (Doctoral dissertation, University of Aberdeen). Retrieved from https://abdn.primo.exlibrisgroup.com/discovery/delivery/44ABE_INST:44ABE_VU1/12174185050005941

Crowther, D. S. (2003). *Key choral concepts: Teaching techniques & tools to help your choir sound great!* Horizon Publishers.

Dahl, D. B., Scott, D., Roosen, C., Magnusson, A., & Swinton, J. (2019). xtable: Export tables to latex or html [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=xtable (R package version 1.8-4)

Day, T. (2018). *I saw eternity the other night: King's College, Cambridge, and an English singing style.* Allen Lane.

Deme, A. (2014). Intelligibility of sung vowels: The effect of consonantal context and the onset of voicing. *Journal of Voice, 28*, 523.e19–523.e25. DOI: https://doi.org/10.1016/j.jvoice.2014.01.003

Demorest, S. M., & Clements, A. (2007). Factors influencing the pitch-matching of junior high boys. *Journal of Research in Music Education, 55*(3), 190–203. DOI: https://doi.org/10.1177/002242940705500302

Dodsworth, R. (2013). Retreat from the Southern Vowel Shift in Raleigh, NC: Social factors. *University of Pennsylvania Working Papers in Linguistics, 19*(5). Retrieved from https://repository.upenn.edu/pwpl/vol19/iss2/5

Dromey, C., Heaton, E., & Hopkin, J. A. (2011). The acoustic effects of vowel equalization training in singers. *Journal of Voice, 25*, 678–682. DOI: https://doi.org/10.1016/j.jvoice.2010.09.003

Fabricius, A. H. (2007). Variation and change in the trap and strut vowels of RP: a real time comparison of five acoustic data sets. *Journal of the International Phonetic Association, 37*(3), 293–320. DOI: https://doi.org/10.1017/S002510030700312X

Ferragne, E., & Pellegrino, F. (2010). Formant frequencies of vowels in 13 accents of the British Isles. *Journal of the International Phonetic Association, 40*(1), 1–34. DOI: https://doi.org/10.1017/S0025100309990247

Friedrichs, D., Maurer, D., Suter, H., & Dellwo, V. (2015). Vowel identification at high fundamental frequencies in minimal pairs. In *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, Scotland.* DOI: https://doi.org/10.5167/uzh-112398

Fromont, R. (2019). Forced alignment of different language varieties using LaBB-CAT. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia.* Australasian Speech Science and Technology Association, Inc. Retrieved from https://assta.org/proceedings/ICPhS2019Microsite/

Fromont, R., & Hay, J. (2012). LaBB-CAT: An annotation store. In *Proceedings of Australasian Language Technology Association workshop* (pp. 113–117). Retrieved from http://hdl.handle.net/10092/15624

Gelman, A. (2020). Prior choice recommendations. Retrieved from https://github.com/stan-dev/stan/wiki/Prior-Choice-Recommendations (Date accessed: 03/03/2021)

Gibson, A. (2019). *Sociophonetics of popular music: Insights from corpus analysis and speech perception experiments* (Unpublished doctoral dissertation). University of Canterbury.

Gibson, A., & Bell, A. (2012). Popular music singing as referee design. In J. A. Cutillas-Espinosa & J. M. Hernández-Campoy (Eds.), *Style-shifting in public: New perspectives on stylistic variation* (pp. 139–164). John Benjamins. DOI: https://doi.org/10.1075/silv.9.08gib

Glasgow Phoenix Choir. (n.d.). *Glasgow Phoenix Choir.* Retrieved from http://www.phoenixchoir.org/ (accessed: 11/04/2022)

Grace, H. (1925). The Glasgow Orpheus Choir. *The Musical Times, 66,* 401–405. DOI: https://doi.org/10.2307/912988

Gregg, J. W., & Scherer, R. C. (2005). Vowel intelligibility in classical singing. *Journal of Voice, 20,* 198–210. DOI: https://doi.org/10.1016/j.jvoice.2005.01.007

Grell, A., Sundberg, J., Ternström, S., Ptok, M., & Altenmüller, E. (2009). Rapid pitch correction in choir singers. *The Journal of the Acoustical Society of America, 126*(1), 407–413. DOI: https://doi.org/10.1121/1.3147508

Harrington, J., Palethorpe, S., & Watson, C. (2000). Monophthongal vowel changes in Received Pronunciation: an acoustic analysis of the Queen's Christmas broadcasts. *Journal of the International Phonetic Association, 30,* 63–78. DOI: https://doi.org/10.1017/S0025100300006666

Hawkins, S., & Midgley, J. (2005). Formant frequencies of RP monophthongs in four age groups of speakers. *Journal of the International Phonetic Association, 35*(2), 183–199. DOI: https://doi.org/10.1017/S0025100305002124

Hillenbrand, J. M., Clark, M. J., & Nearey, T. M. (2001). Effects of consonant environment on vowel formant patterns. *The Journal of the Acoustical Society of America, 109*(2), 748–763. DOI: https://doi.org/10.1121/1.1337959

Hollien, H., Mendes-Schwartz, A. R., & Nielsen, K. (2000). Perceptual confusions of high-pitched sung vowels. *Journal of Voice, 14*(2), 287–298. DOI: https://doi.org/10.1016/S0892-1997(00)80038-7

Hollins, L., & Vango, S. (2022). *How to make your choir sound awesome!* Banks Music Publications.

Hughes, A., Trudgill, P., & Watt, D. (2012). *English accents and dialects: An introduction to social and regional varieties of English in the British Isles* (Fifth Edition ed.). Routledge. DOI: https://doi.org/10.4324/9780203784440

Johnston, A. (2016). *English and German diction for singers: A comparative approach* (Second ed.). Lanham, Maryland: Rowman & Littlefield.

Johnston, P. (1997). Regional variation. In C. Jones (Ed.), *The Edinburgh history of the Scots language* (pp. 378–432). Edinburgh University Press. DOI: https://doi.org/10.1515/9781474410977-013

Johnston, P. A. (1985). The rise and fall of the Morningside/Kelvinside accent. In M. Görlach (Ed.), *Focus on Scotland: Varieties of English around the world* (Vol. 5). John Benjamins Publishing Company. DOI: https://doi.org/10.1075/veaw.g5.04joh

Jones, M. (1989). A study of vocal pitch-matching skills among undergraduate education majors using classroom instruments. *Update: Applications of Research in Music Education, 7*(2), 39–41. DOI: https://doi.org/10.1177/875512338900700213

Karreth, J., Scogin, S., Williams, R., & Beger, A. (2021). BayesPostEst: Generate postestimation quantities for Bayesian MCMC estimation [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=BayesPostEst (R package version 0.3.1)

King, J. B., & Horii, Y. (1993). Vocal matching of frequency modulation in synthesized vowels. *Journal of Voice, 7*, 151–159. DOI: https://doi.org/10.1016/S0892-1997(05)80345-5

Krause, M., & Smith, J. (2017). 'I stole it from a letter, off your tongue it rolled' the performance of dialect in Glasgow's indie music scene. In C. Montgomery & E. Moore (Eds.), *Language and a sense of place: Studies in language and region.* Cambridge University Press.

Labov, W. (1972). *Sociolinguistic patterns.* University of Pennsylvania Press.

Leech-Wilkinson, D. (2009). *The changing sound of music: Approaches to studying recorded musical performances.* London: CHARM. Retrieved from https://www.charm.rhul.ac.uk/studies/chapters/intro.html

Leifeld, P. (2013). texreg: Conversion of statistical model output in R to LATEX and HTML tables. *Journal of Statistical Software, 55*(8), 1–24. DOI: https://doi.org/10.18637/jss.v055.i08

Lenth, R. V. (2021). emmeans: Estimated marginal means, aka least-squares means [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=emmeans (R package version 1.5.4)

Lévêque, Y., Giovanni, A., & Schön, D. (2012). Pitch-matching in poor singers: Human model advantage. *Journal of Voice, 26*(3), 293–298. DOI: https://doi.org/10.1016/j.jvoice.2011.04.001

Lobanov, B. M. (1971). Classification of Russian vowels spoken by different speakers. *The Journal of the Acoustical Society of America, 49*(2B), 606–608. DOI: https://doi.org/10.1121/1.1912396

Marshall, E. J. (2023a). *Do choirs have accents? A sociophonetic investigation of choral sound* (Doctoral dissertation). University of Glasgow. theses.gla.ac.uk/id/eprint/84372

Marshall, E. J. (2023b). O Lo/r/d, open thou ou/r/ lips: Rhoticity in choral singing from Glasgow and Cambridge. In Radek Skarnitzl & Jan Volín (Eds.), *Proceedings of the 20th International Congress of Phonetic Sciences,* Prague (pp. 2184–2188). Guarant International.

Marvin, J. (1991). Choral singing, in tune. *Choral Journal, 32*(5), 27. Retrieved from https://www.proquest.com/docview/1306222128

Morrissey, F. A. (2008). Liverpool to Louisiana in one lyrical line: Style choice in British rock, pop and folk singing. In M. A. Locher & J. Strässler (Eds.), *Standards and norms in the English language.* De Gruyter Mouton. DOI: https://doi.org/10.1515/9783110206982.1.195

Potter, J. (1998). *Vocal authority: Singing style and ideology.* Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511470226

Powell, S. (1991). Choral intonation: More than meets the ear. *Music Educators Journal, 77*(9), 40–43. DOI: https://doi.org/10.2307/3398190

R Core Team. (2021). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from https://www.R-project.org/

Rathcke, T., Stuart-Smith, J., Torsney, B., & Harrington, J. (2017). The beauty in a beast: Minimising the effects of diverse recording quality on vowel formant measurements in sociophonetic real-time studies. *Speech Communication, 86*, 24–41. DOI: https://doi.org/10.1016/j.specom.2016.11.001

Riegle, A. M., & Gerrity, K. W. (2011). The pitch-matching ability of high school choral students. *Update: Applications of Research in Music Education, 30*(1), 10–15. DOI: https://doi.org/10.1177/8755123311418618

Roberton, H. S., & Roberton, K. (Eds.). (1963). *Orpheus with his lute: A Glasgow Orpheus Choir anthology*. Pergamon Press.

Sagrans, J. (2016). *Early music and the Choir of King's College, Cambridge, 1958 to 2015* (Unpublished doctoral dissertation). Schulich School of Music, McGill University, Montreal.

Shekar, P., & Fujioka, T. (2014). The effects of timbre and musical training on vocal pitch-matching accuracy. In *Proceedings of the International Conference on Music Perception and Cognition, San Francisco, California.* DOI: https://doi.org/10.13140/2.1.2038.1445

Simpson, P. (1999). Language, culture and identity: With (another) look at accents in pop and rock singing. *Multilingua – Journal of Cross-Cultural and Interlanguage Communication, 18*, 343–367. DOI: https://doi.org/10.1515/mult.1999.18.4.343

Smith, L. A., & Scott, B. L. (1980). Increasing the intelligibility of sung vowels. *Journal of the Acoustical Society of America, 67*, 1795–1797. DOI: https://doi.org/10.1121/1.384308

Sóskuthy, M., & Stuart-Smith, J. (2020). Voice quality and coda /r/ in Glasgow English in the early 20th century. *Language, Variation and Change, 32*, 133–157. DOI: https://doi.org/10.1017/S0954394520000071

Stuart-Smith, J. (1999). Glasgow: Accent and voice quality. In G. Docherty & P. Foulkes (Eds.), *Urban voices: Accent studies in the British Isles* (pp. 203–222). Taylor & Francis.

Stuart-Smith, J. (2003). The phonology of modern urban Scots. In J. Corbett, J. D. McClure, & J. Stuart-Smith (Eds.), *The Edinburgh companion to Scots* (pp. 110–137). Edinburgh University Press. DOI: https://doi.org/10.1515/9781474421591-010

Sundberg, J. (1974). Articulatory interpretation of the "singing formant". *The Journal of the Acoustical Society of America, 55*(4), 838–844. DOI: https://doi.org/10.1121/1.1914609

Sundberg, J. (1987). *The science of the singing voice.* Northern Illinois University Press.

Sundberg, J., & Ternström, S. (1986). Acoustic comparison of voice use in solo and choir singing. *Journal of the Acoustical Society of America, 79*(6), 1975–1981. DOI: https://doi.org/10.1121/1.393205

Trudgill, P. (1983/1997). Acts of conflicting identity: The sociolinguistics of British pop-song pronunciation. In N. Coupland & A. Jaworski (Eds.), *Sociolinguistics: A reader* (pp. 251–265). Macmillan Education UK. (Originally published in Trudgill, P. (1983), *On Dialect: Social and Geographical Perspectives, Blackwell*). DOI: https://doi.org/10.1007/978-1-349-25582-5_21

Wells, J. C. (1982a). *Accents of English: An introduction* (Vol. 1). Cambridge University Press.

Wells, J. C. (1982b). *Accents of English: The British Isles* (Vol. 2). Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511611766

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. Retrieved from https://ggplot2.tidyverse.org. DOI: https://doi.org/10.1007/978-3-319-24277-4

Wilson, G. A. F. (2014). *The sociolinguistics of singing: Dialect and style in classical choral singing in Trinidad* (Unpublished doctoral dissertation). Westfälischen Wilhelms-Universität, Münster.

Wilson, G. A. F. (2017). Conflicting language ideologies in choral singing in Trinidad. *Language & Communication*, *52*, 19–30. DOI: https://doi.org/10.1016/j.langcom.2016.08.003

Yang, J. H. (2018). 'I want to be new and different. anything I'm not.' Accent-mixing in singing. *Australian Journal of Linguistics*, *38*(2), 183–204. DOI: https://doi.org/10.1080/07268602.2018.1400501

Young, S., Evermann, G., Gales, M., Hain, T., Kershaw, D., Liu, X. A., … Woodland, P. (2009). The HTK book (for HTK version 3.4) [Computer software manual]. Cambridge University Engineering Department. Retrieved from https://htk.eng.cam.ac.uk/docs/docs.shtml