

IJRCS

International Journal of Research in Choral Singing

The Scientific Research Journal of the American Choral Directors Association

International Journal of Research in Choral Singing
(2026) Vol. 14 154-178

“O–hel–pu–slor”: Realization of word-final /d/ in British classical choral singing

Edward J. Marshall

Abstract

Previous research has found that the accent of British classical choral singing is based on a non-regional prestige accent, Southern Standard British English, in terms of front vowels. Evidence shows that local accents can also form part of the choral accent, in terms of consonants. However, choral singing as an art form is more than a direct translation of a spoken accent to a sung medium—there are also musical aesthetic variables that create the overall character of choral sound. One such variable in classical choral singing is word-final /d/. This article asks how the realization of word-final /d/ changes over time in recordings of the Choir of King’s College, Cambridge, and the Glasgow Orpheus and Glasgow Phoenix choirs. Here, I present evidence of a reduction in voicing over time in Glasgow, and increased affrication over time at King’s, which may relate to the directorial influence of Stephen Cleobury, to singing with orchestral accompaniment, or to an ongoing change in spoken English. These findings demonstrate how choral diction norms and choir director stylistic influence may shape the sound of British classical choirs over time.

Keywords: *choral diction, British choral style, choir director, phonetic variation, consonant realization*

Corresponding author: Edward_Marshall1@outlook.com

British classical choral singing has changed over the twentieth century in the pronunciation of front vowels (as in the words *kit*, *dress* and *trap*) which have lowered following an established pattern in a prestigious British spoken accent (Marshall et al., 2024). Crucially, this vowel lowering has been found in regions with differing spoken accents, Glasgow English and Southern Standard British English, suggesting there are shared features of British classical choral singing based on a spoken standard accent. In contrast, regional accents have been found to influence the choral accent. E.g., whether /r/ is pronounced following a vowel (e.g. *car*) has been found to vary in choral singing by spoken accent region (Marshall, 2023; 2026). The role of choral directors such as David Willcocks, in shaping a choir's sound is thought to have driven a substantial change in the vowel sounds of the Choir of King's College, Cambridge (Marshall et al., 2024). Here, I present the first quantitative sociolinguistic study to focus on a musical–phonetic variable that is not thought to vary based on the linguistic system of the participants but is instead an iconic aesthetic feature of choral style: word-final /d/. Analyzing this aesthetic variable will provide insight into how diction choices in classical choirs makes the choir's sound evolve over time.

Most previous sociolinguistic research on singing has focused on linguistically motivated variables, despite an acknowledgment that singing is not speech and is “subject to different parameters” (Morrissey, 2008, p. 213). In choral literature, word-final consonants, particularly stops, have been identified as presenting difficulties for audibility (e.g., Robertson & Robertson, 1963, p. 50). For choral singers and practitioners, word-final consonants are not only linguistic units, but part of the ensemble's style. Using quantitative sociophonetic methods (see Kendall et al., 2023), this study investigates the role of the choir director in shaping the realization of word-final /d/ in choral singing.

In this article, I use standard linguistic conventions such that slashes e.g. /d/ refer to phonemes (abstract sound categories); square brackets e.g. [d] refer to specific phonetic realizations. The word “realization” refers to a way that a phoneme is produced in a given instance. Word examples are given in italics.

Stops in Choral Literature

Many choral directors that have written about diction in choral singing in English have noted variation in the realization of word-final, and particularly, phrase-final consonants. E.g., the conductor of the Glasgow Orpheus Choir, Hugh S. Robertson writes that singers frequently produce phrases like:

“‘O–hel–pu–slor’ under the mistaken idea that they are singing – ‘O help us, Lord!’ [...] Why are singers so afraid to *close* words which have consonant endings? Not one singer in a hundred can sing ‘and’. It scares them to death” (Robertson & Robertson, 1963, p. 50).

The quotation from Robertson & Robertson demonstrates the salience of word-final con-

sonants. They imply that singers are naturally inclined not to “close” syllables, e.g. articulate word-final consonants. As a remedy to this, many directors explicitly promote the use of “shadow vowels” (epenthetic vowels), as in *lord* realized as [lɔːd̥] or [lɔːd̥ʰ], to increase the audibility of final consonants (e.g., Coward, 1914; Cappadonia, 1962; Moore, 1972; Decker, 1977; LaBouff, 2008; Neuen, 2020). Neuen (2020) gives justification for the use of epenthetic vowels following final consonants, arguing that the “preceding vowel and/or the accompaniment usually covers them up” (Neuen, 2020, p. 78). In short, vowels are more sonorous than consonants, so adding a following short vowel can highlight the consonant. While shadow vowel realizations are recommended for all voiced English word-final consonant phonemes, there is a particular focus on word-final and phrase-final /d/, which need extra attention to ensure audibility in singing (Henkel, 1965). Johnston (2016) writes that one of the pitfalls for learning to sing in English is the voicing of final voiced consonants, noting that it is difficult for German native speakers to distinguish between *hand* [hænd] and *hand* [hant]. Crucially, Johnston also notes that this tendency is “rampant among English native speakers as well” (2016, p.139).

Previous Sociolinguistic Research on Stops in Singing

Previous sociolinguistic research on singing has been restricted largely to popular styles (Morrissey, 2008). One of the most salient features of British Englishes, both Anglo and Scottish, is the prevalence of glottal stops as a possible realization of /t/. Readers may recognize this from the stereotyping of British accents in the phrase *bottle of water* as [bɒtəl əv 'wɔːtə]. T-glottaling has been investigated in sociolinguistic studies of popular singing (e.g., Trudgill, 1983/1997; Simpson, 1999; Beal, 2009; Westphal & Jansen, 2021). Generally, it has been interpreted that if bands seek to align themselves with the popular standard, they produce tap [ɾ] variants of /t/ (e.g., *water* as in General American ['wɑːrə]), as in the USA-5 model (Trudgill, 1983/1997; Simpson, 1999; Gibson, 2019; Gibson, 2024). If bands seek to project a regional identity, then they diverge from the norms of the style by selecting regional variants; For many British bands, this meant employing the glottal stop [ʔ] variant, e.g.: Ian Drury (Trudgill, 1983/1997), Dire Straits (Simpson, 1999), Arctic Monkeys (Beal, 2009), and most recently, Stormzy (Westphal & Jansen, 2021). In the sociolinguistic singing literature, much less attention has been paid to the realization of /d/.

Phonetic Background: Stops in English

Stops are a family of consonants produced with three phases: closure, hold, and release. The release phase is what we are most interested in as different types of release can affect the audibility of a stop consonant in choral singing. Stops require a complete closure of the vocal tract, allowing air pressure to build up. After some time, the pressure is released as a small burst or puff of air (Catford, 2002; Ogden, 2009). The location of the closure can vary, i.e., distinguishing /p, t, k/. At each location of closure, the consonant produced can be “voiced” or “voiceless” (e.g., /t/ vs /d/). In this article, I focus on /d/ (the voiced alveolar

plosive); however, note that category may be realized as phonetically fully voiced, partially voiced, or voiceless.

In English, stops can appear in word-initial, word-medial or word-final positions (e.g., *day* /deɪ/, *addle* /adl/, or *kid* /kɪd/). Both word position and surrounding phonetic context can affect the realization of the stop. Word-final /d/ is also theoretically motivated from a linguistic perspective, as word-final stops may be more salient than in other contexts and therefore “it is more likely that speakers would implement a stylistic use of release in this position” (Davidson, 2011, p. 1050).

In addition to voicing and location, there are several features of stops which can vary, including aspects relating to the release (or burst). Word-final stops can be articulated or deleted (not produced at all). Stops that are articulated can be released or unreleased, e.g., *good* /gʊd/ can be released [gʊd^h], or unreleased [gʊd̚] where there is a closure and hold phase, but no burst. Unreleased realizations are very common in English and particularly in word-final contexts (Lisker & Abramson, 1967), or when preceding another stop, e.g., *good boy* produced as [gʊd̚bɔɪ]. Davidson notes that “increased prevalence of stop release may be a hallmark of a more careful or formal speech register, such as read speech” (2011, p. 1056). Perhaps we might expect to find more released stops in classical choral singing, which is analogous to formal speech. Realizations of /d/ that are articulated (not deleted) and released can differ in various aspects of the release.

When the tongue parts from the roof of the mouth to produce the burst, it can do so swiftly, or more gradually, producing a transitional [s] sound, or [z] in the case of a voiced stop. This is called affrication. /s/ has been called “a horror” in choral singing (Decker, 1977, p. 6) due to how audible it is. So, it may follow that affricated variants of stops like /t/ or /d/ are more audible because of that /s/-like quality. In English it is common for /t/ to be affricated [tʰ] and affrication can also occur for /d/ ([dʰ]) (Lindsey, 2019, p. 56). In a study of /t/ affrication, Buizza and Plug (2012) find that affricated realizations of /t/ are significantly longer and louder than other realizations of /t/. Their conclusion that stop-affrication is a form of strengthening complements the singing literature which recommends “wet t/d” (affrication) for audibility. E.g., LaBouff writes:

“With thick accompaniment or orchestration, use a wet t/d to project. A wet t/d has a forward-placed vowel released with it. Release final t’s with a whispered [ɪ] vowel following it. A final d is released with an [ɪ] vowel following it. E.g., *night* [naɪt⁽ⁱ⁾]. This final wet t sounds similar to the percussion instrument, the high hat” (2008, p. 124).

As we have seen, word-final /d/ has several possible realizations in spoken English. In choral singing, affricated and shadow vowel variants are thought to be most audible and could aid intelligibility in typical choral acoustics or when singing with accompaniment. Increasing affrication may be part of a larger change in progress in global English and goes together with singing pedagogy.

Research Questions

So far, I have outlined various ways in which stop voicing and burst have been reported to vary, both in speech and in singing. To examine how ensemble diction practices shape choral sound over time, this study addresses the following research questions: How is word-final /d/ realized in British classical choral singing, and how do patterns of voicing and affrication differ across ensembles and phonetic contexts? How have realizations of word-final /d/ changed over time? And to what extent do choir directors contribute to observed changes in the realization of word-final /d/, beyond what might be expected from spoken-language change alone?

Method

Sample

The corpus analyzed here features commercially-released recordings of the Choir of King’s College, Cambridge (1945–2019) and the Glasgow Orpheus and Phoenix choirs (1925–2016). The combined corpora feature over 25 hours of unaccompanied choral singing in English. The Choir of King’s College, Cambridge was selected as the prototypical collegiate choir due to its prolific recording history and the perceived social salience of the King’s sound in the British choral context (Potter, 1998; Day, 2018). The corpus includes extracts from 317 tracks from 50 albums. Table 1 summarizes the King’s choir directors over the recorded period.

Table 1
Summary of the choir directors of the Choir of King’s College, Cambridge

Choir	Dates active	Director	Coding
The Choir of King’s College, Cambridge	1929-1940	Boris Ord	
	1940-1945	Harold Darke	
	1945-1958	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	

The Glasgow choirs were selected to provide a contrasting dialect region to King's, and because the Glasgow Orpheus Choir's prestige and broadcast history, together with the continuity of the Glasgow Phoenix Choir, offer a consistent record of performance practice. The corpus includes extracts from 178 tracks (songs) from 28 albums. Table 2 summarizes the Glasgow choir directors over the recorded period.

Table 2

Summary of the choir directors of the Glasgow Orpheus and Glasgow Phoenix choirs

Choir	Dates active	Director	Coding
Glasgow Orpheus	1901–1951	Hugh S. Robertson	HSR 1925–1951
Glasgow Phoenix	1955–1983	Peter Mooney	PM 1959–1975
	1984–1986	John Cranston	
	1986–1991	Peter S. Shand	
	1991–2016	Marilyn J. Smith	MJS 1987–2016
	2018–present	Cameron Murdoch	

Together, these two corpora allow for a comparison between a socially prestigious collegiate choir and a long-standing regional amateur tradition, providing insight into how stylistic norms and directorial influence shape consonant realization across contrasting choirs. The choirs also differ in membership: King's is comprised of adult male singers and boy choristers, whereas the Glasgow choirs are adult mixed-voice ensembles.

An electronic corpus was constructed in LaBB-CAT, a browser-based linguistics research tool that stores audio, text transcripts, and other annotations (Fromont & Hay, 2012; Fromont, 2019). The lyrics and audio (.wav) files were automatically aligned using the Hidden Markov Toolkit (HTK) forced aligner (Young et al., 2006). HTK is used to automatically time-align audio recordings of speech with their corresponding text transcripts. The resulting time-alignments of choral singing and lyrics were hand-corrected. Each token (individual instance) of word-final /d/ was extracted from the corpus and auditorily coded.

Auditory Coding Protocol

Analyzing recordings of choral singing presents the “joint speech/joint singing” problem (Cummins, 2018; Marshall, 2026). Choirs contain multiple people singing at the same time, however, no matter how good the singers are, coordination will not always be perfect. Where a token was articulated, but there were multiple different phonetic realizations produced simultaneously, I endeavored to code the majority percept, that is, the variant that I perceived as most prominent and/or that most people were producing.

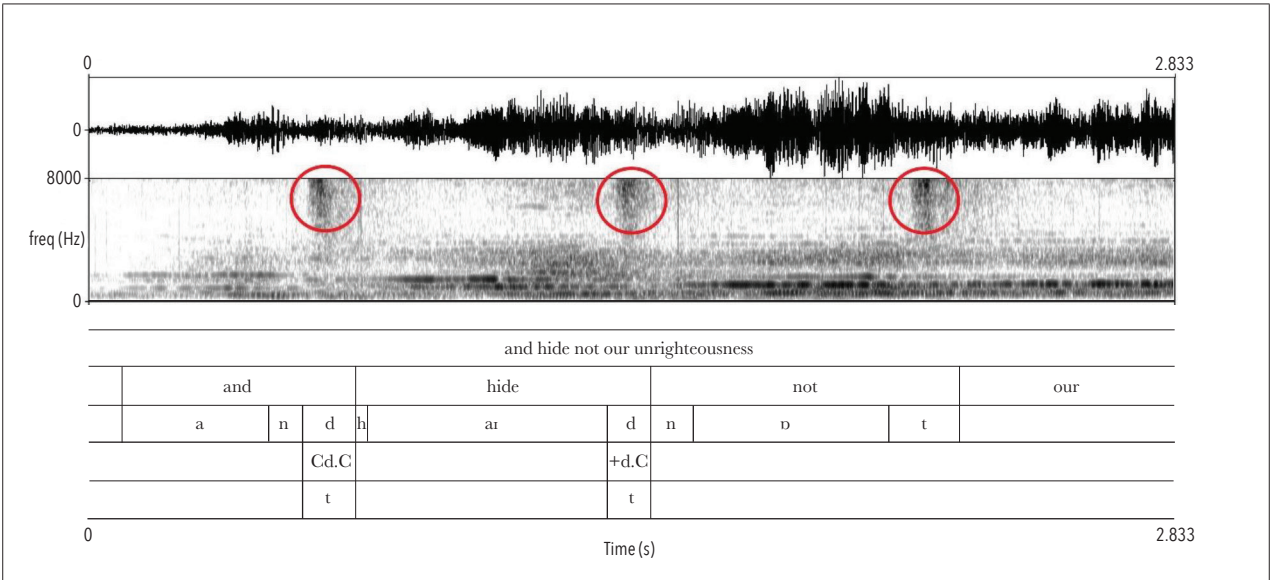
I systematically listened to the audio files in conjunction with a visual inspection of the waveform spectrogram and coded the realization in a Praat textgrid (Boersma & Weeninck,

2018). Praat is a computer programme designed to analyze recordings of speech. Textgrids are Praat text files that are used to store time-aligned transcripts and annotations (see Fig. 1). I carried out a phonetic transcription using the International Phonetic Alphabet, which resulted in the following fine-grained phonetic categories:

- [d^ɹ] – hyper-voiced (with shadow vowel/epenthetic schwa)
- [d^{zə}] – hyper-voiced + affricated
- [d⁼] – “normal” (audible burst, usually elided with following syllable)
- [d^z] – “normal” + affricated
- [d[̚]] – unreleased (no audible burst)
- [t⁼] – voiceless unaffricated
- [t^s] – hyper-voiceless (affricated)
- [Ø] – zero (no audible percept of /d/)

Differentiating affricated from aspirated realizations auditorily presented a challenge, so for the purposes of this study, these instances were combined and will be referred to only as Affricated. The Affricated tokens are visible in spectrograms as clouds of high-frequency energy, which resemble fricatives as in *sue* or *shoe*. This reflects the fact that affrication produces a noisier, more audible release. As shown in Figure 1, affricated realizations of /d/ (*and*, *hide*) can be indistinguishable from typical realizations of /t/ (*not*), producing the same high-frequency noise described by LaBouff (2008) as resembling a high hat cymbal.

Figure 1
Waveform spectrogram of affricated realizations of /d/ with /t/ for comparison.



Note. Example from “Hide not Thou Thy face from us” from Evensong for Ash Wednesday (1964). Recording by the Choir of King’s College Cambridge, directed by David Willcocks. This figure shows a waveform spectrogram annotated with a textgrid. Here, the textgrid has five tiers of annotations. The first three tiers show the transcript, word, and sound that are extracted from the corpus. The lower two tiers show /d/ context, and /d/ realization that were manually coded.

A full breakdown of the realizations by Corpus is found in Table 3, showing that the most frequent variant in Glasgow was [d^ə], and the most frequent variant in Cambridge was [d^z], both voiced. Overall affricated realizations are more common in the recordings of King's.

Table 3

Distribution of variants by Corpus

Corpus	Variant	N	(%)
Glasgow	[d ^ə]	207	18.8%
Glasgow	[d ^{zə}]	34	3.0%
Glasgow	[d ⁼]	277	25.2%
Glasgow	[d ^z]	134	12.2%
Glasgow	[d ^ɹ]	183	16.6%
Glasgow	[t ⁼]	134	12.2%
Glasgow	[t ^s]	17	1.5%
Glasgow	[Ø]	115	10.5%
King's	[d ^ə]	168	6.8%
King's	[d ^{zə}]	103	4.1%
King's	[d ⁼]	385	15.1%
King's	[d ^z]	535	21.5%
King's	[d ^ɹ]	484	19.5%
King's	[t ⁼]	196	7.9%
King's	[t ^s]	356	14.3%
King's	[Ø]	258	10.4%

Table 4

Distribution of Voicing by Corpus

Corpus	N	Voiced	Voiceless	% Voiced
Glasgow	986	835	151	85%
King's	2227	1675	552	75%

Note. The category Voiced includes the variants: [d^ə], [d^{zə}], [d⁼], [d^z], [d^ɹ]; the category Voiceless includes variants: [t⁼], [t^s].

In this article, I present two binary analyses of these data. First, a binary analysis of Voicing—whether a token is ‘voiced’ or ‘voiceless’. Second, a binary analysis of Affrication—whether a token is affricated or not. The category Voiced groups the variants: [d^ɹ], [d^{ɹʷ}], [d[̃]], [d^ʰ], [d[̚]]; the category Voiceless groups variants: [t[̚]], [t^s]. A breakdown of the number of Voiced tokens by Corpus is given in Table 4 on the previous page.

The category Affricated groups the variants: [d^ʒ], [d^{ʒ̥}], and [t^s] (i.e., releases with an added ‘s-like’ quality that may increase clarity in ensemble singing); the category Not Affricated groups the variants: [d^ɹ], [d[̃]], [d[̚]], and [t[̚]]. A breakdown of the number of Affricated tokens by Corpus is provided in Table 5.

Table 5
Distribution of Affrication by Corpus

Corpus	N	Affricated	Unaffricated	% Affricated
Glasgow	986	185	801	18.8%
King's	2227	994	1233	44.6%

Note. The category Affricated includes the variants: [d^ʒ], [d^{ʒ̥}], and [t^s]; the category Not Affricated includes variants: [d^ɹ], [d[̃]], [d[̚]], and [t[̚]].

Word-Final /d/ Phonetic Contexts

For the analyses reported here, /d/ is always postvocalic, i.e., following a vowel (e.g., *good*). The preceding vowels are almost always stressed (vowels in choral singing are more often stressed than in speech), so the preceding context is controlled for and not explicitly included in the model. However, any effect of stress will also be partially accounted for in the modeling by the varying effects structures, which include intercepts for Word, and by Corpus or by Time/Director slopes for Word.

The following phonetic context has been shown to affect the realization of word-final /d/. E.g., a following pause increases the likelihood that /d/ is unreleased (Davidson, 2011). Therefore, following context is explicitly included in the model. Optimally, I would include following segment (with a level for each type of following consonant, vowel, or pause); however, due to the number of tokens, I decided to collapse to a three-level factor Context with the levels: “d.V” with following vowel (e.g., *good apple, and again*); “d.C” following consonant (e.g., *good boy, and saw*); and, “d#” following pause. In practical terms, the effect of context will help to explain why some /d/ sounds are lost in choral performance. When followed by another consonant, or occurring before a pause, /d/ is more prone to disappearing. For choir directors, this suggests that phrase-final and pre-consonantal instances of /d/ may require deliberate shaping to ensure clarity, which might take the form of adding time, or specifying shadow vowels, or encouraging affricated release in the rehearsal. Understanding these contexts will allow directors to anticipate where diction may need reinforcement.

Exclusions

Davidson (2011) excludes the function word *and* due to extreme reduction e.g., being produced only as [n] (e.g., *rock ‘n’ roll*) or even schwa [ə]. However, in the data reported here, *and* is not reduced to the extent found in speech – apart from in some rare cases. Following Gibson and Bell (2012), I decided to include function words where they were not reduced. Additionally, Davidson (2016) excludes /d/ before /t, d, ð/ “since /d/ was never released in this environment and therefore could not be distinguished from the following sound” (p. 36). However, in these corpora, the realization is variable; as Neuen (2020) notes, double consonants are usually articulated in classical choral singing.

Any tokens where multiple different words were being sung at once were excluded. Any instances of audible noise that were visible in spectrograms, such as clicks that were artifacts of the recording or digitization process, were discarded. 3,755 tokens of word-final /d/ were coded. 169 tokens were excluded, leaving 3,586 tokens. Of these, 373 tokens were deletions. These were also removed, leaving 3,213 tokens for the analyses presented here.

Predictions

In the singing literature, affricated realizations of stops are prescribed when the singer is required to project in a large acoustic, or over orchestral accompaniment (LaBouff, 2008). Musicologists have noted a change to the sound of King’s under the direction of David Willcocks, which he himself attributed to more frequent singing with orchestral accompaniment (Day, 2002; Williams, 2019). Perhaps, therefore, we could expect more affricated tokens under David Willcocks than under his predecessor Boris Ord. Similarly, the Orpheus and Phoenix choirs rarely sang with accompaniment, and when they did, mainly with piano and organ. Consequently, we might expect fewer affricated stops in the Glasgow corpus than in the King’s corpus. An increase in affrication over time in the King’s data would also be supported by change over time in speech from Received Pronunciation to Southern Standard British English (Lindsey, 2019).

Based on Day’s (2018) comments about King’s pronunciation under David Willcocks (“I know thett my Redeemer liveth, ent thett he shell stent”, p. 261), we might expect more voiceless realizations of /d/. Drawing from spoken realization, we might expect recordings of the Glasgow choirs to have more voiced variants overall (Wells, 1982; Masuya, 1988; Catford, 2002), though these differences to Southern Standard British English may have lessened over time (Stuart-Smith et al., 2015).

Modeling

For the Voicing analysis, the models reported investigate the effect of region Corpus and Context, or Time/Director and Context, on whether a token of word-final /d/ is Voiced or Voiceless. For these models, Voiced = 1 and Voiceless = 0, such that the greater the proportion/percentage, the more likely a token is to be Voiced. Similarly, for Affrication, the

models reported investigate the effect of Corpus and Context, or Time/Director and Context on whether a token of word-final /d/ is Affricated or Not Affricated. For these models, Affricated = 1 and Not Affricated = 0, such that the greater the proportion/percentage, the more likely a token is to be Affricated. Please note that zero (deletion) is excluded from both analyses. The data are analyzed with Bayesian binomial logistic mixed models using brms (Bürkener, 2018) [version 2.18.0] in R (R Core Team, 2021) [version 4.1.2]. For all models reported here, the estimates are on the log-odds scale. A negative logit difference means the variant is less likely than the average; a positive value means it is more likely. In the text, these will be converted to probabilities. Further details can be found in Appendix 1.

Results

Voicing

Glasgow Diachronic Voicing Model

The model reported in this section asks whether there is evidence of change over time in the Glasgow corpus. I.e., do tokens of word-final /d/ in recordings made under different directors have different likelihoods of being Voiced or Voiceless? Also, does voicing vary by Context (following vowel, consonant, or pause), and is there an interaction between Time/Director and Context?

Table 6

Glasgow /d/ voicing posterior summary.

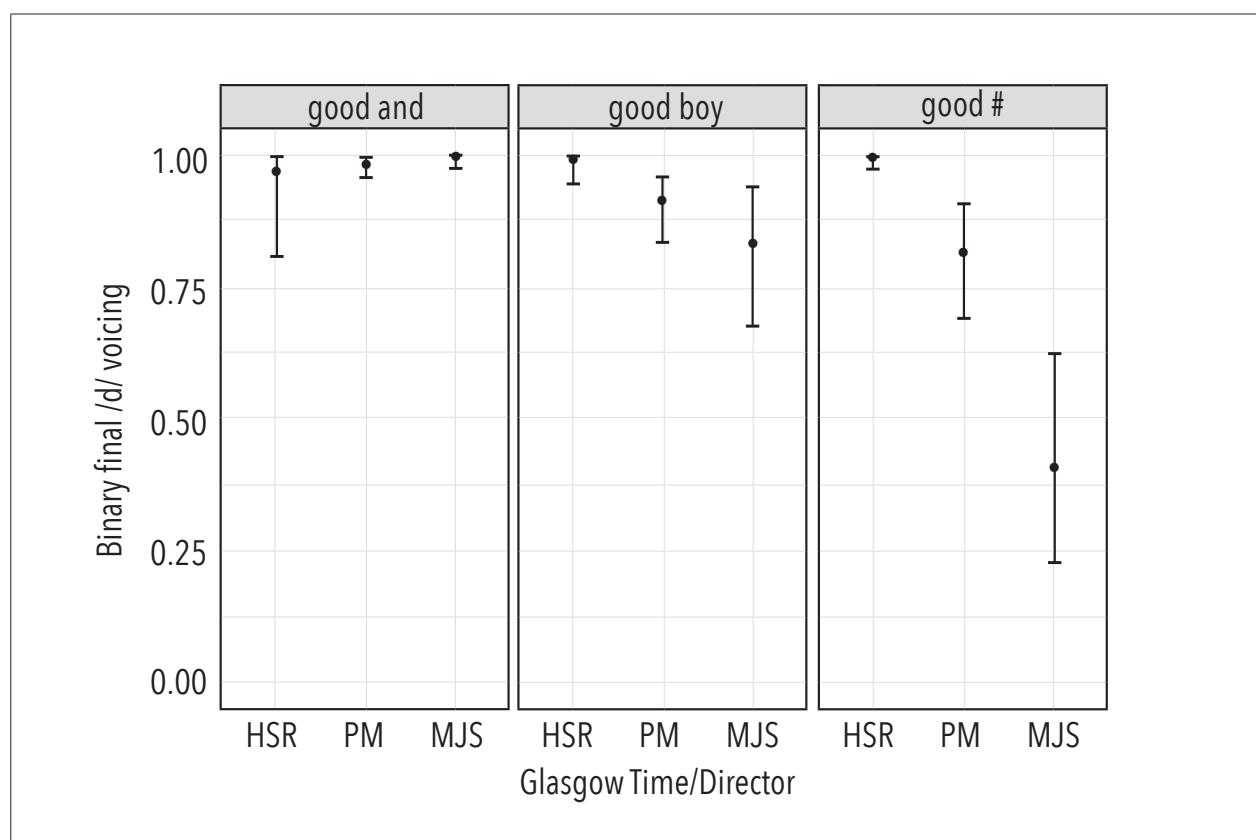
Parameter	Estimate	95% CI	
Intercept	3.9	2.85	5.35
Time/DirectorPM (1959–1975)	-1.16	-2.57	-0.09
Time/DirectorMJS (1987–2016)	-0.74	-2.25	1.23
Context_d.C	-0.87	-2.2	0.24
Context_dp	-0.78	-2.18	1.19
DirectorPM:Context_d.C	0.5	-0.64	1.86
DirectorMJS:Context_d.C	-0.65	-2.69	0.86
DirectorPM:Context_dp	-0.48	-2.42	0.98
DirectorMJS:Context_dp	-2.86	-5.34	-1.19

Note. Point estimate displayed: median. Results are given on the log odds ratio (not the response) scale. Red type indicates 0 outside 95% credible interval.

Note in Table 6 on the previous page that the Intercept is strongly positive with a predicted probability of being Voiced of 0.98 (logit: 3.90, CI [2.85; 5.35]). There is little evidence supporting change over time overall. However, Time/Director PM is less likely to produce voicing with a predicted probability of 0.93 (logit difference -1.16, CI [-2.57; -0.09]). There is an interaction of Time/Director by Context as tokens produced under Marilyn J. Smith are substantially less likely to be Voiced in prepausal Contexts (predicted probability 0.4; logit difference -2.86, CI [-5.34; -1.19]). The interaction is visualized in Figure 2.

Figure 2

Glasgow corpus word-final /d/ voicing model interaction of Time/Director by Context.



Note. HSR = Hugh S. Robertson (1925–1951); PM = Peter Mooney (1959–1975); MJS = Marilyn J. Smith (1987–2016).

These findings suggest that the Glasgow choirs have generally maintained voicing of word-final /d/, except in phrase-final positions where the consonant is most vulnerable to devoicing. The decrease in voicing under Marilyn J. Smith, particularly before pauses, indicates a stylistic shift toward lighter or less emphatic final consonants over time.

King's Diachronic Voicing Model

The following model asks whether there is evidence of change over time in the King's corpus. I.e., do tokens of word-final /d/ in recordings made under different directors have different likelihoods of being Voiced or Voiceless? Also, does voicing vary by Context (fol-

lowing vowel, consonant, or pause), and is there an interaction between Time/Director and Context?

Table 7

King's /d/ voicing posterior summary.

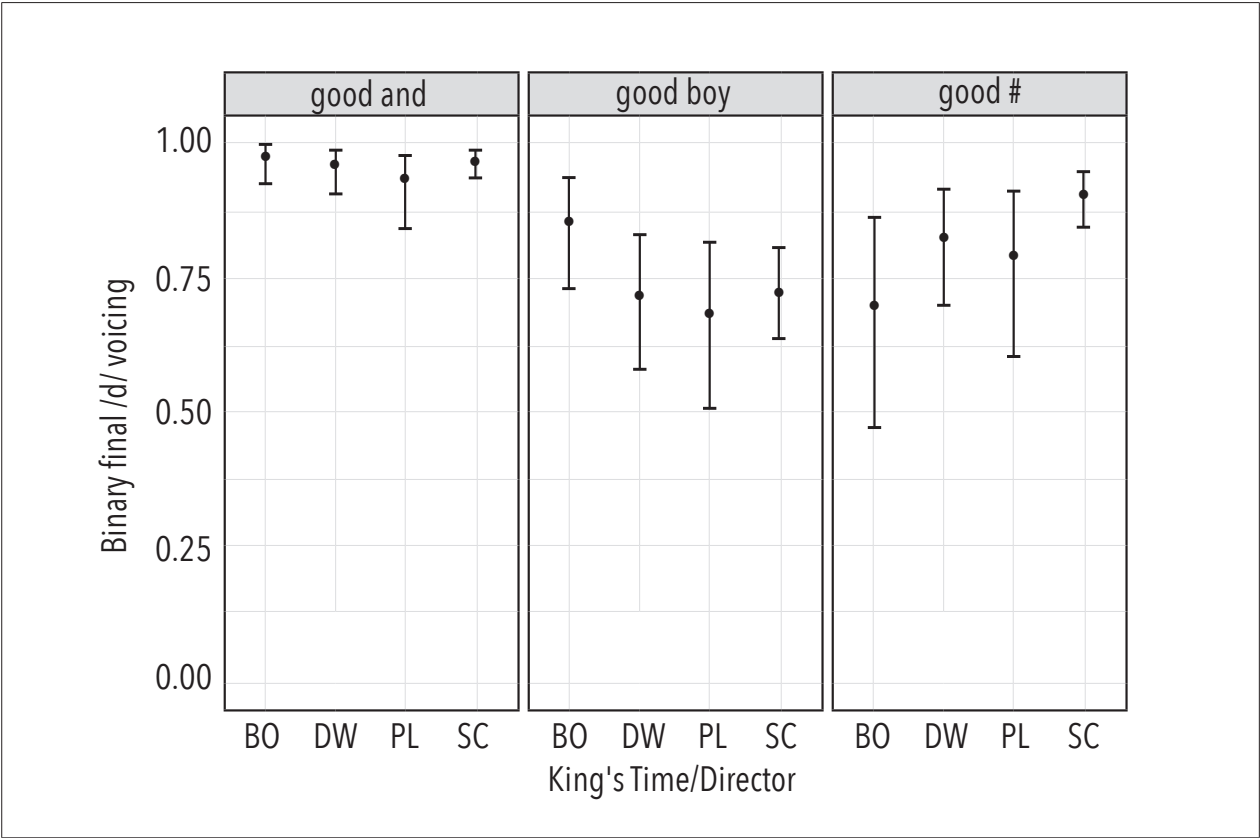
Parameter	Estimate	95% CI	
Intercept	1.93	1.59	2.29
Time/DirectorDW (1959–1974)	-0.07	-0.60	0.50
Time/DirectorPL (1976–1982)	-0.35	-0.98	0.25
Time/DirectorSC (1984–2019)	0.25	-0.20	0.71
Context_d.C	-0.83	-1.03	-0.63
Context_dp	-0.45	-0.71	-0.19
DirectorDW:Context_d.C	-0.11	-0.45	0.21
DirectorPL:Context_d.C	0.01	-0.35	0.36
DirectorSC:Context_d.C	-0.39	-0.67	-0.10
DirectorDW:Context_dp	0.13	-0.28	0.55
DirectorPL:Context_dp	0.21	-0.27	0.68
DirectorSC:Context_dp	0.49	0.12	0.87

Note. Point estimate displayed: median. Results are given on the log odds ratio (not the response) scale. Red type indicates 0 outside 95% credible interval.

As seen in Table 7, the Intercept for the King's binary Voicing model is strongly positive with a predicted probability of /d/ tokens being Voiced of 0.87 (logit: 1.93, CI [1.59; 2.29]). The model does not provide evidence for a main effect of Time/Director. There is evidence for a main effect of Context with preconsonantal tokens less likely to be Voiced (predicted probability 0.75; logit difference -0.83, CI [-1.03; -0.63]). Prepausal tokens are also less likely to be Voiced (predicted probability 0.81; logit difference -0.45, CI [-0.71; -0.19]). There is an interaction of Time/Director by Context with tokens of final /d/ produced under Stephen

Cleobury less likely to be voiced in preconsonantal contexts (predicted probability 0.72; logit difference -0.39, CI [-0.67; -0.10]) and more likely to be Voiced before a pause (predicted probability 0.90; logit difference 0.49, CI [0.12; 0.87]). The interaction of Time/Director by Context is visualized in Figure 3.

Figure 3
King’s corpus word-final /d/ voicing model interaction of Time/Director by Context.



Note. BO = Boris Ord (1945–1958); DW = David Willcocks (1959–1974); PL = Philip Ledger (1976–1982); SC = Stephen Cleobury (1984–2019).

The King’s results show that voicing patterns vary subtly by director, with Stephen Cleobury favoring increased voicing before pauses but less voicing before consonants. This aligns with a more “legato-driven” approach to text adopted under Cleobury, compared to the clipped sound preferred by David Willcocks.

Affrication

Glasgow Diachronic Affrication Model

The following model asks whether the probability of word-final /d/ Affrication changes over time in the Glasgow corpus.

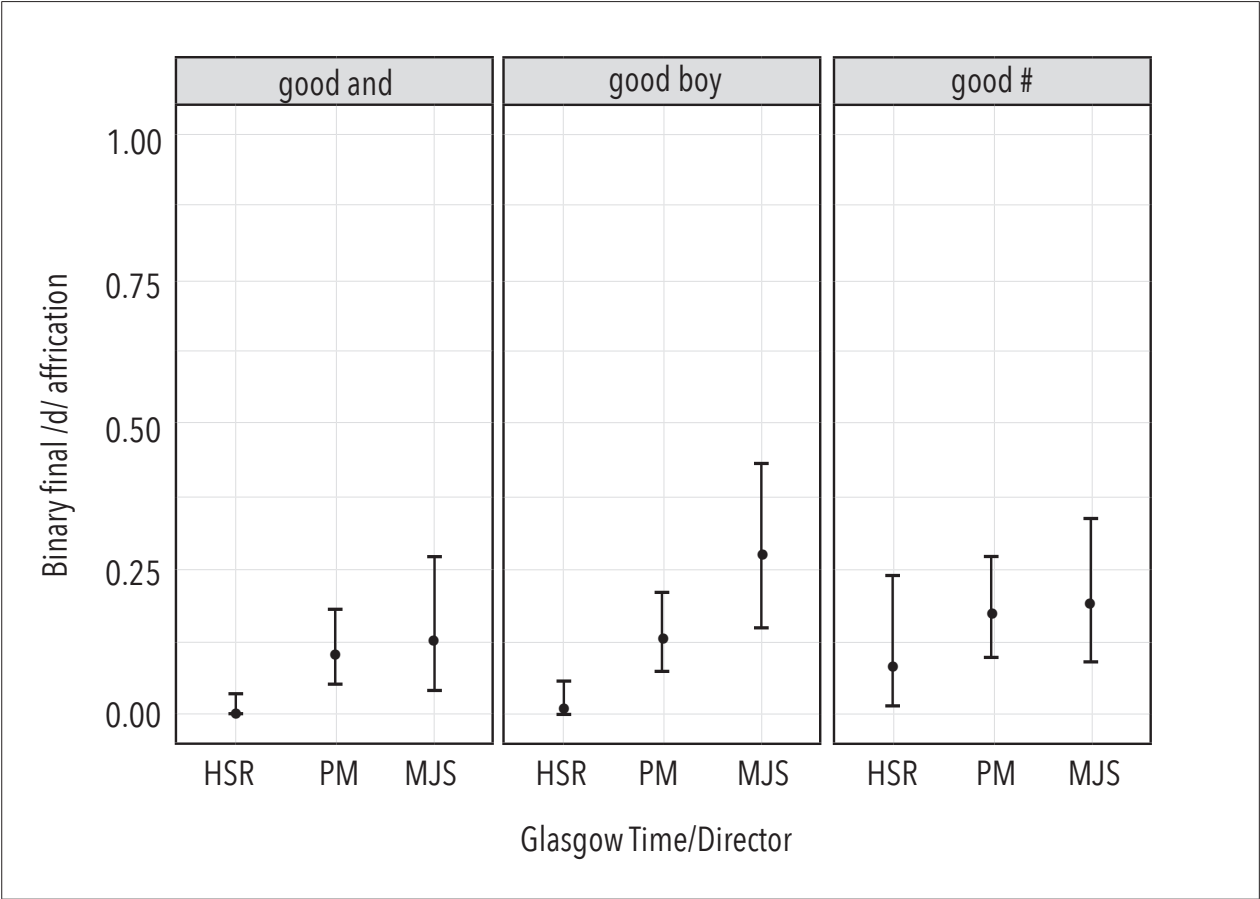
Table 8*Glasgow /d/ affrication posterior summary.*

Parameter	Estimate	95% CI	
Intercept	-2.90	-4.23	-2.07
Time/DirectorPM (1959–1975)	1.05	0.21	2.40
Time/DirectorMJS (1987–2016)	1.45	0.59	2.81
Context_d.C	0.37	-0.56	1.72
Context_dp	1.10	0.33	2.42
DirectorPM:Context_d.C	-0.39	-1.75	0.57
DirectorMJS:Context_d.C	0.11	-1.26	1.09
DirectorPM:Context_dp	-0.81	-2.14	-0.00
DirectorMJS:Context_dp	-1.10	-2.45	-0.25

Note. Point estimate displayed: median. Results are given on the log odds ratio (not the response) scale. Red type indicates 0 outside 95% credible interval.

As seen in Table 8, the model intercept is strongly negative with a predicted probability of Affrication of 0.05 (logit: -2.9, CI [-4.23; -2.07]). The model supports a main effect of Time/Director whereby /d/ becomes more likely to be Affricated over time; Tokens produced under Peter Mooney (1959–1975) are more likely to be Affricated than the grand mean with a predicted probability of 0.135 (logit difference 1.05, CI [0.21; 2.40]). The same pattern is strengthened under Marilyn J. Smith (1987–2016) with a predicted probability of 0.19; logit difference 1.45, CI [0.59; 2.81]. The model also supports a main effect of Context with prepausal tokens more likely to be Affricated (predicted probability 0.14; logit difference 1.10, CI [0.33; 2.42]). There is an interaction of Time/Director by Context. Due to the positive main effect of Time/Director, there is a slight negative adjustment for prepausal tokens for both Peter Mooney (predicted probability 0.17; logit difference -0.81, CI [-2.14; -0.00]) and Marilyn J. Smith (predicted probability 0.19; logit difference -1.10, CI [-2.45; -0.25]). The interaction of Time/Director by Context is visualized in Figure 4 on the next page.

Figure 4
Glasgow corpus final /d/ affrication model interaction of Time/Director by Context.



Note. HSR = Hugh S. Robertson (1925–1951); PM = Peter Mooney (1959–1975); MJS = Marilyn J. Smith (1987–2016).

Affrication remains rare in the Glasgow choirs, but the gradual increase over time suggests that affrication may be spreading more widely within British classical choral singing. The slightly higher rates before pauses may indicate that singers tend to employ affrication in these contexts to enhance consonant audibility.

King’s Diachronic Affrication Model

The final model in this section asks whether the probability of affrication of word-final /d/ changes over time in the King’s corpus.

As seen in Table 9 in the next page, the Intercept for the model is negative with a predicted probability of Affrication of 0.29 when all other effects are kept constant. The model supports a main effect of Time/Director with tokens of /d/ produced under Philip Ledger (1976–1982) more likely to be Affricated (predicted probability 0.47; logit difference 0.78, CI [0.23; 1.33]). There is also a main effect of Context with preconsonantal tokens more likely to be Affricated (predicted probability 0.41; logit difference 0.54, CI [0.37; 0.72]). There is an interaction of Time/Director by Context with tokens more likely to be Affricated in pre-

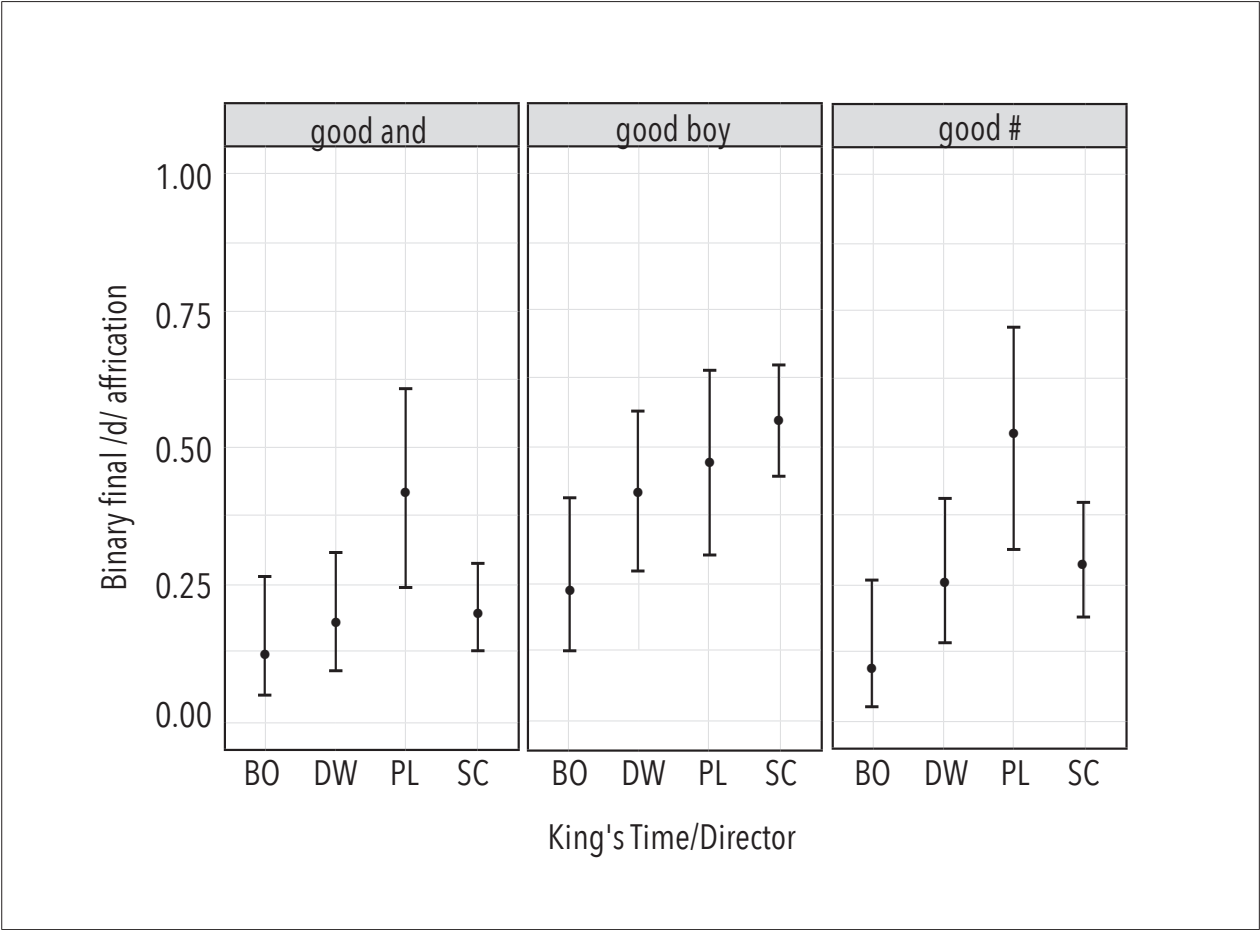
Table 9*King's /d/ affrication posterior summary.*

Parameter	Estimate	95% CI	
Intercept	-0.89	-1.23	-0.56
Time/DirectorDW (1959–1974)	-0.08	-0.59	0.42
Time/DirectorPL (1976–1982)	0.78	0.23	1.33
Time/DirectorSC (1984–2019)	0.19	-0.21	0.60
Context_d.C	0.54	0.37	0.72
Context_dp	-0.13	-0.40	0.12
DirectorDW:Context_d.C	0.08	-0.19	0.36
DirectorPL:Context_d.C	-0.54	-0.85	-0.24
DirectorSC:Context_d.C	0.37	0.13	0.59
DirectorDW:Context_dp	0.04	-0.33	0.43
DirectorPL:Context_dp	0.36	-0.06	0.43
DirectorSC:Context_dp	-0.07	-0.40	0.26

Note. Point estimate displayed: median. Results are given on the log odds ratio (not the response) scale. Red type indicates 0 outside 95% credible interval.

consonantal contexts under the direction of Stephen Cleobury (predicted probability 0.55; logit difference 0.37, CI [0.13; 0.59]). In contrast, preconsonantal tokens are less likely to be Affricated under Philip Ledger compared to the mean effect of Philip Ledger (predicted probability 0.47; logit difference -0.54, CI [-0.85; -0.24]). The interaction of Time/Director by Context for the King's /d/ Affrication model is visualized in Figure 5 on the next page.

Figure 5
King’s word-final /d/ affrication model interaction of Time/Director by Context.



Note. BO = Boris Ord (1945–1958); DW = David Willcocks (1959–1974); PL = Philip Ledger (1976–1982); SC = Stephen Cleobury (1984–2019).

The King’s data show a marked rise in affrication over time, especially under Philip Ledger and in preconsonantal contexts. This aligns with the widely noted “crispness” of the King’s sound and suggests that affrication became an intentional stylistic feature.

Discussion

This study investigated the realization of word-final /d/ in British classical choral singing from Glasgow and Cambridge. 3,213 tokens of /d/ were auditorily coded and analyzed with respect to Voicing and Affrication using Bayesian binomial mixed models.

In the voicing analysis /d/ variants were collapsed into categories Voiced ([d^ɹ], [d^z], [d[̥]], [d[̌]], and [d[̎]]) and Voiceless ([t[̥]], [t[̌]]). Overall, the predicted probability of tokens to be perceived as Voiced by the researcher was 0.90. /d/ tokens were less likely to be Voiced before a consonant or before a pause. Preconsonantal tokens were more likely to be Voiced in the Glasgow corpus than in the King’s corpus. Conversely, prepausal tokens were more likely to be Voiceless in the Glasgow corpus than in the King’s corpus. In the Glasgow corpus, I find

that /d/ is significantly more likely to be Voiceless in prepausal contexts for Time/Director Marilyn J. Smith (1987–2016), and there is a trend of Voicing decreasing over time in both prepausal and preconsonantal contexts. In the King's corpus, I find that /d/ is less likely to be Voiced in preconsonantal contexts for Time/Director Stephen Cleobury (1984–2019) and more likely to be voiced in prepausal contexts. For King's, there is a trend of increasing Voicelessness in preconsonantal contexts and increasing Voicing in prepausal contexts.

In the affrication analysis /d/ variants were collapsed into the categories Affricated ([d^{zə}], [d^z], [t^s]) and Unaffricated ([d^ə], [d^ː], [d̪], [t^ː]). The overall predicted probability of Affrication was 0.20. Affrication was quite rare overall in the Glasgow corpus, but there was a significant increase over time. There was also an increase of Affrication over time in the King's corpus in preconsonantal Contexts, increasing from approximately 20% for Boris Ord (1945–1958) to around 50% for Stephen Cleobury (1984–2019). For King's, word-final /d/ is least likely to be Affricated in the context of a following vowel and most likely to be Affricated when followed by a consonant.

As we saw in the voicing models, prepausal tokens have become significantly less likely to be voiced in the Glasgow recordings, and conversely significantly more likely to be voiced in the King's recordings. This interaction is unlikely to be driven by changes in spoken language, but rather a change in the choir's aesthetic, either driven by the director, or by the choir members.

The realization of word-final /d/ was selected for investigation as it is salient to the overall aesthetic of classical choirs. Voiceless realizations are more common before a consonant, as expected based on findings in speech that adjacent fricatives or plosives are likely to promote devoicing regardless of whether they are voiced or voiceless (Davidson, 2016). Likewise, it is consistent with phonetic findings that tokens of /d/ before a pause are less likely to be Voiced than tokens followed by a vowel (Davidson, 2016). However, Davidson finds that for stops in prepausal contexts, the predicted probability of partial or full voicing was 0.57 (compared to voiceless). While we cannot directly compare the auditory coding of this work and the acoustic analysis conducted by Davidson, it seems that there is a tendency for more frequent voicing of phonologically voiced stops in prepausal contexts in singing than in speech. This would fit with the view that phonation during singing is prioritized more highly than in speech. I.e., voiced consonants can be voiced at the pitch of the adjacent vowel to carry the tune (Morrissey, 2008; Johnston, 2016, p. 104). While this tendency is unsurprising, it needs to be empirically tested by recording individual singers producing voiced stops and then using a similar methodology to measure the proportion and shape of voicing distribution during the stop closure. The greater frequency of voicing I find here compared to speech more generally may also reflect the greater attention-to-speech (Labov, 1979) being paid in classical choral singing.

The increase in affrication over time in both the Glasgow and King's corpora may mimic increases in affrication in Southern Standard British English (Lindsey, 2019) and Standard Scottish English (Chirrey, 1999) speech and may reflect a broader change in English globally (Dodsworth & Mielke, 2025). It is also possible that singing pedagogy has played a role

in the increase of affrication in these data, as affrication is recommended to help boost the audibility of word-final /d/ when singing with orchestral accompaniments (LaBouff, 2008).

For the Glasgow Orpheus choir under Hugh S. Robertson (1925–1951), /d/ is almost always Voiced in all contexts. The increase in the use of Voiceless variants in the recordings of the Glasgow Phoenix choir could reflect a realization of word-final /d/ more closely aligned to the variant distribution found in spoken Scottish English, which includes (voiceless) ejective allophones of voiced stops (Gordeeva & Scobbie, 2013). Contrary to the attention-to-speech model interpretation of the categorical realization of rhoticity (Marshall, 2026), phrase-final /d/ is not consistently Voiced in the Glasgow choir singing. This shows that attention to speech does not apply to all variables uniformly, and there may be some variables which are more salient to a particular choir's style than others.

At King's, the increase in the use of Voiced variants in prepausal contexts over time, and the increase in the consistency of Voiced variants under Stephen Cleobury (1984–2019) particularly, demonstrates how prepausal Voiced /d/ has become a feature of the style of King's, and perhaps also the style of elite British classical choral singing more widely.

Affricated variants of /d/ have been incorporated into the King's style, particularly in preconsonantal contexts. While there has been an increase in the proportion of affricated variants over time in the Glasgow corpus, it remains to be seen whether affricated variants of /d/ have become a feature of a broader British classical choral aesthetic. Affricated realizations of /d/ were substantially less frequent in the Glasgow corpus compared to the King's corpus. The relatively lower rate of affrication in Glasgow could relate to affrication being associated with Southern Standard British English, or perhaps it aligns with our musically informed predictions that choirs/singers need to affricate /d/ when accompanied by an orchestra or for singing in the acoustic of the chapel of King's College, Cambridge (LaBouff, 2008; Williams, 2019). The choirs were not accompanied in the recordings analyzed here; however, it is possible that the habits formed when singing with organ and orchestra—the latter becoming more frequent after the arrival of David Willcocks—were carried over into choral *cappella* singing, as suggested by Day (2018, p. 173).

These stylistic differences should also be interpreted with consideration of the membership of the different choirs. King's is comprised of boy choristers that sing the top line, and adult male lay clerks and choral scholars that sing the alto, tenor, and bass parts. In contrast, the Glasgow choirs are both adult mixed-voice choirs. These differences could potentially influence articulatory precision. However, I consider it unlikely that these differences in gender and age between the corpora will have affected the findings of change over time within-choir, as the demographics of the respective choirs have stayed relatively constant over time.

The choirs also differ in their contexts and rehearsal processes, which may shape their approaches to diction. King's is a collegiate choir with a long-established history, including daily services during term time, regular international tours, commercial recordings, and national broadcasts. The Glasgow Orpheus Choir, by contrast, was an ensemble whose reputation was built through extensive broadcasting and recording in the interwar period and touring internationally before it became commonplace (Robertson & Robertson, 1963, p. 5).

Under Hugh S. Robertson, the Orpheus maintained a demanding rehearsal schedule and held annual reauditions of the whole choir. The Glasgow Phoenix choir was founded by former members of the Orpheus in a similar ethos. The culture and membership of the choir changed over time and it no longer has competitive auditions. However, the choir has recorded an album nearly every year since 1959. These cultural and organizational differences may influence expectations around diction. So, it is possible that the reduction in consonant voicing before pauses observed in the Glasgow data may, in part, reflect these broader contextual factors.

The findings of this study suggest that word-final /d/ functions as an ensemble-level stylistic resource and highlights how diction contributes to an identifiable British classical choral sound. In terms of the role of the choir director, the observed changes in affrication and voicing over time may be cultivated implicitly or in the absence of explicit instruction. These results invite directors and choral educators to reflect on how diction norms are communicated, whether deliberately or otherwise. More broadly, this study highlights the value of focusing on fine-grained phonetic detail as a meaningful aspect of choral style. Future research could analyze the impact of musical dynamic on the realization of word-final /d/. E.g., Johnston writes that ‘A plosive [d] will not poke out of the melodic line if the surrounding phonemes are similar in intensity. A consonant must match its environment’ (2016, p. 110). Does the dynamic of a sung passage affect how choirs collectively realize word-final and phrase-final /d/?

The findings presented here highlight how word-final /d/ is sensitive to both phonetic context and director and they offer several points of practical relevance for choir leaders. First, directors may wish to give particular attention to phrase-final and preconsonantal /d/, which are the contexts most prone to devoicing across the recordings examined here. In these positions, shadow vowels or affrication may improve intelligibility. Second, the diachronic patterns observed at King’s and in Glasgow show that diction norms are not fixed but evolve over time; directors therefore play a central role in shaping how consonants function within their ensemble’s sound. Finally, the increase of affrication, especially in the King’s recordings, suggests that this technique may be an effective tool for choirs when projecting text in resonant acoustics or over accompaniment. Moreover, through the popularity and broadcast frequency of King’s (Day, 2018), affrication and shadow vowel realizations may be becoming part of the wider British classical choral style.

Conclusion

This article demonstrates how the realization of word-final /d/ contributes to British classical choral sound and accent. The Choir of King’s College, Cambridge is more likely to use shadow vowel realizations before a pause, and affricated variants in preconsonantal contexts. In contrast, in the Glasgow choirs, voiceless realizations have become more common over time. When followed by a vowel, the realization of word-final /d/ may be affected by whether the choir director and singers prioritize intelligibility or legato line. It is unclear

whether the realization of word-final /d/ is influenced by a choral target accent, but it seems likely that shadow vowels and affrication have become a style marker of liturgical music and British classical choral singing more generally. By treating word-final consonants as ensemble-level stylistic variables, this study demonstrates how phonetic detail is a productive area for future choral research bridging performance practice, pedagogy, and empirical analysis.

Acknowledgments

This research was supported by the UK Arts and Humanities Research Council (project code 2284740).

Ethical Approval

As I am working with commercially released recordings and/or public broadcasts, this study did not require ethical approval.

Data Availability

The datasets generated and analyzed during the current study are available on the Open Science Framework <https://osf.io/8xgwk/>. If you would like access to the corpus for checking, please contact the author.

References

- 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(3), 233–240. <https://doi.org/10.1177/007542420934001>
- Boersma, P., & Weenink, D. (2018). Praat: Doing phonetics by computer [computer program]. (Version 6.0.43). Retrieved from: <https://praat.org/>.
- Buizza, E. & Plug, L. (2012). Lenition, fortition and the status of plosive affrication: the case of spontaneous RP English /t/. *Phonology*, 29(1), 1–38. <https://doi.org/10.1017/S0952675712000024>
- Bürkener, P. C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, 10, 395–411. <https://doi.org/10.32614/RJ-2018-017>
- Cappadonia, A. C. (1961). The importance of the clarity and production of consonants. *Choral Journal*, 2(2), 16. Retrieved from <https://www.jstor.org/stable/23541659>
- Catford, J. C. (2002). *A practical introduction to phonetics*. Oxford University Press, 2nd ed.
- Chirrey, D. (1999). Edinburgh: descriptive material. In Foulkes, P. & Docherty, G. J. (Eds.) *Urban voices: Accent studies in the British Isles* (pp. 223–229). Arnold.
- Coward, H. (1914). *Choral technique and interpretation*. Novello.
- Cummins, F. (2018). *The ground from which we speak: Joint speech and the collective subject*. Cambridge Scholars Publishing.

- Davidson, L. (2011). Characteristics of stop releases in American English spontaneous speech. *Speech Communication*, 53(8), 1042–1058.
<https://doi.org/10.1016/j.specom.2011.05.010>
- Davidson, L. (2016). Variability in the implementation of voicing in American English obstruents. *Journal of Phonetics*, 54(1), 35–50.
<https://doi.org/10.1016/j.wocn.2015.09.003>
- Day, T. (2002). *A century of recorded music: Listening to musical history*. Yale University Press.
- Day, T. (2018). *I saw eternity the other night: King's College, Cambridge, and an English singing style*. Allen Lane.
- Decker, W. (1977). Can your choristers sing? Diction. *Choral Journal*, 17(9), 5. Retrieved from <https://www.proquest.com/scholarly-journals/can-your-choristers-sing-diction/docview/1306215764/se-2>
- Dodsworth, R., & Mielke, J. (2025). /t d/ Releases are strengthening among White speakers: Evidence from a large-scale acoustic study of English in Raleigh. *Language Variation and Change*, 37(1), 1–30. doi:10.1017/S095439452500002X
- 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.
- Gibson, A. (2019). Sociophonetics of popular music: Insights from corpus analysis and speech perception experiments. [Doctoral dissertation, University of Canterbury (New Zealand)]. UC Repository. <http://dx.doi.org/10.26021/4007>.
- Gibson, A. (2024). Pop song English as a supralocal norm. *Language in Society* 53(3). 471–98. <https://doi.org/10.1017/S0047404523000131>.
- Gordeeva, O. B. & Scobbie, J. M. (2013). A phonetically versatile contrast: Pulmonic and glottalic voicelessness in Scottish English obstruents and voice quality. *Journal of the International Phonetic Association*, 43(3), 249–271.
 doi:10.1017/S0025100313000200
- Henkel, G. J. (1965). What is good choral diction? *Choral Journal*, 5(6). Retrieved from <https://www.proquest.com/scholarly-journals/what-is-good-choral-diction/docview/1306214513/se-2>
- Johnston, A. (2016). *English and German diction for singers: A comparative approach*. Rowman & Littlefield, 2nd ed.
- Kendall, T., Phrao, N., Stuart-Smith, J., & Vaughn, C. (2023). Advancements of phonetics in the 21st century: Theoretical issues in sociophonetics. *Journal of Phonetics*, 98, 1–37. <https://doi.org/10.1016/j.wocn.2023.101226>
- LaBouff, K. (2008). *Singing and communicating in English: A singer's guide to English diction*. Oxford University Press.
- Lindsey, G. (2019). *English after RP: Standard British pronunciation today*. Palgrave Macmillan.

- Lisker, L. & Abramson, A. S. (1967). Some effects of context on voice onset time in English stops. *Language and Speech*, 10(1), 1–28.
<https://doi.org/10.1177/002383096701000101>
- Masuya, Y. (1988). Voice onset time of the syllable-initial /p/, /t/ and /k/ followed by an accented vowel in Lowland Scottish English. In *Onseigaku to oninron: shuyo ronko [Phonetics and phonology: selected papers]*, (pp.139–172).
- Marshall, E. J., Stuart-Smith, J., Butt, J. & Dean, T., (2024). “Variation and change over time in British choral singing (1925–2019)”, *Laboratory Phonology*, 15(1), 1–39.
 doi: <https://doi.org/10.16995/labphon.10125>
- Marshall, E. J. (2023). “O Lo/r/d, open thou ou/r/ lips: rhoticity in choral singing from Glasgow and Cambridge” In R. Skarnitzl, & J. Volín (Eds.) *Proceedings of the 20th International Congress of Phonetic Sciences* (pp. 2184–2188). Guarant International.
- Marshall, E. J. (2026). “Sweet singing in the choi/r/”: How spoken dialect affects choral phonology in Glasgow and Cambridge. *Journal of English Linguistics* (Accepted for publication). <https://eprints.gla.ac.uk/381623/>.
- Moore, R. R. (1972). Toward a better concept of choral diction through the International Phonetic Alphabet. *Choral Journal*, 13(1). Retrieved from <https://www.jstor.org/stable/23543846>
- 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.
- Neuen, D. (2020). Rehearsal break: Diction and textual artistry: Successful communication when singing in English. *Choral Journal*, 60(10), 77–85.
- Ogden, R. (2009). *An introduction to English phonetics*. Edinburgh University Press.
- R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Version 4.2.2, retrieved from <https://www.R-project.org/>.
- Roberton, H. S., & Roberton K. (1963). *Orpheus with his lute: A Glasgow Orpheus Choir anthology*. Pergamon Press.
- 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(4), 343–367. <https://doi.org/10.1515/mult.1999.18.4.343>.
- Stuart-Smith, J., Sonderegger, M., Rathcke, T., & Macdonald, R. (2015). The private life of stops: VOT in a real-time corpus of spontaneous Glaswegian. *Laboratory Phonology*, 6(3–4), 505–549. <https://doi.org/10.1515/lp-2015-0015>.
- Trudgill, P. (1983/1997). *On dialect: Social and geographical perspectives*. Blackwell.
- Young, S., Evermann, G., Gales, M., Hain, T., Kershaw, D., Liu, X. A., Moore, G., Odell, J., Ollason, D., Povey, D., Valtchev, V., & Woodland, P. (2009). The HTK book (for HTK Version 3.4). Retrieved from: <https://htk.eng.cam.ac.uk/docs/docs.shtml>. Accessed: 26/05/24.

Appendix 1

The model structure for each corpus was:

Voicing or Affrication $\sim$ Time/Director + Context + Time/Director: Context + (1 | Album)
+ (1 | Song) + (1 | Time/Director: Word)

All factor variables are sum coded, meaning each level reported in the model summary is compared to the grand mean for that factor. The priors used were Cauchy (0, 10) for Intercepts, Cauchy (0, 2.5) for varying effects, and normal (0, 2.5) for fixed effects. Model chains were visually inspected for convergence, R_{hat} was 1 for all coefficients, and there were no divergent transitions after warmup. The minimum effective sample size for all coefficients was greater than $100 \times$ the number of chains. I was satisfied that the models converged successfully and that the posterior summaries were amenable to interpretation. Posterior predictive checks for evaluating model fit to the observed data can be found in supplementary materials.
