

Diction, expressivity and conductor's choices in choral works sung in German

Caiti Hauck

University of Bern, Switzerland

ABSTRACT: Musical performance requires several choices by the performer. In choral performance, choices include aspects related to text diction, such as the duration and dynamics of vowels and consonants. These choices may be targeted, for example, at text comprehensibility, although they may also contribute to achieving an expressive performance. This article examines conductors' performance choices regarding text diction in choral works sung in German. It aims to identify possible uses of diction — mainly by means of manipulations in the duration of consonants — as an expressive device and to discuss the effects of such use on performance expressivity. Methods consisted, firstly, of analyses of writings on choral conducting, which aimed to identify suggestions concerning the expressive use of diction. Secondly, six choral conductors were interviewed. Thirdly, choral works by Johann Sebastian Bach and Franz Schubert were used as case studies, which consisted of score analyses and, principally, of recording analyses done with the software Sonic Visualiser. Finally, the data collected were compared and discussed. Results show that some choral conductors point to the expressive potential of text diction, especially of consonants. Results also show that, in some recordings, important words are highlighted by means of manipulating the duration and dynamics of consonants. Such manipulations can be understood as expressive gestures that not only reinforce the meaning of a piece, but also create different meanings. When these manipulations are frequent, results suggest that diction can play a central role in the expressivity of choral performance.

KEY WORDS: choral, diction, expressivity, performance, recordings

Musical performance requires choices (Rink, 2002). Far from being the reproduction of a score, performance involves several decisions related, for example, to elements that musical notation cannot specify (e.g. articulation, timbre, variations in tempo or dynamics), to the performer's understanding of a composition, to the technical demands of a work, or to the acoustics of the hall where the performance occurs.

In choral performances, these choices are, at least in part, the conductor's responsibility.¹ He or she may make decisions concerning tempo and dynamics, or also choral timbre, balance, and blend. Besides that, the choral conductor may make choices related to text diction, partly because Western musical notation is not precise about the duration and the dynamics of vowels and consonants (for instance, how long and how loud each diphthong vowel or a voiced consonant should sound), as well as exactly when an initial or a final consonant, for example, should be articulated.

Choral conducting handbooks generally discuss text diction in terms of uniformity of pronunciation and clearness of enunciation, occasionally describing the pronunciation of different languages (e.g. Bastian & Fischer, 2006, pp. 258-261; Emmons & Chase, 2006, pp. 60-100; Garretson, 1998, pp. 90-99, 106-107; Hammar, 1984, pp. 77-78; Kaplan, 1985, pp. 57-76; Pöhlmann, 1981, pp. 127-128; Smith & Sataloff, 2006, pp. 197-201), or also in terms of the relationship of vowels and consonants with choral blend and/or with a legato line (e.g. Garretson, 1998, pp. 91-92, 96, 99-105, 217; Hammar, 1984, pp. 75-76, 190-191; Kaplan, 1985, p. 60). But despite the necessity for making choices concerning the aspects of diction that musical notation does not specify, only a few writings on choral conducting are explicit about possible durations, dynamics, and/or instants of articulation of vowels and consonants (e.g. Blocker, 2004, pp. 101-103; Kaplan, 1985, pp. 61-68; Thomas, 1979, pp. 90-98), and even rarer are those illustrating how these choices may contribute to an expressive performance (Ehmann & Haasemann, 1990, p. 159; Halsey, 2011, pp. 210-211).

In recent decades, research on expressivity in musical performance has increased. Focusing on expressivity in the singing voice, authors such as Jansens, Bloothoof, and Krom (1997), Leech-Wilkinson (2009), Rapoport (1996), Scherer, Sundberg, Tamarit, and Salomão (2015), Siegwart and Scherer (1995), Sundberg (2000), Sundberg, Iwarsson, and Hagegård (1994), Sundberg, Lã, and Himonides (2013), and Timmers (2007) have studied expressivity in solo singing, analysing elements such as vocal timbre, intonation, vibrato, variations in tempo or dynamics, and others. Of these authors, Leech-Wilkinson (2009), Sundberg (2000), and Sundberg et al. (1994) discuss the expressive potential of text diction. Despite increased interest in understanding performance expressivity, little is known about the expressive use of diction specifically in choral performances, and about the effects of such performance choices.

This article aims to identify choral conductors' performance choices regarding text diction — mainly choices concerning the duration of consonants — and to discuss the effects of these choices on the expressivity of choral performance.² Choices related to diction are relevant in all languages, and some procedures may be similar in many languages (for example the

¹ I appreciate that singers may influence conductor's choices, for instance, by means of the way they sing and by proposing ways of performing a piece that the conductor may not have thought about. Nonetheless, this article focuses on the conductor's point of view, not least because studying singers' perceptions would require a different object of study and different methods of analysis.

² I appreciate that a performance may have (different) effects on (different) audiences, singers, or conductors. In this article, I focus on the effects I noticed when listening to the recordings I analysed.

anticipation of initial consonants, as discussed later). However this discussion will focus on the German language, in view of the large choral repertoire sung in German and also of the specific knowledge necessary to discuss the diction of each language.

The research used a mixed methods approach. Firstly, it consisted of bibliographical studies, in which writings on choral conducting were analysed, aiming to identify suggestions related to the expressive use of diction. Secondly, the conductors Georg Christoph Biller, Timothy Brown, John Butt, Stephen Cleobury, Martin Ennis, and Peter Neumann were interviewed; questions related to their approach to diction in choral rehearsals and performances, and about possible uses of diction as an expressive device. Thirdly, choral works by Johann Sebastian Bach and Franz Schubert were used as case studies. These consisted of analyses of the score and, principally, of analyses of recordings made using the software Sonic Visualiser (Cannam, Landone, & Sandler, 2010) and also by listening critically. Analyses of recordings aimed to identify manipulations in the duration and dynamics (i.e. loudness) of consonants and, occasionally, of vowels.³

The article begins with a short description of research on the expressivity of text diction in sung performances. This is followed by an account of the main data collected in the bibliographical studies, in the interviews, and in the case studies. Finally, I compare the data collected, discussing the implications of these results for expressivity in choral performance.

Expressivity in vocal performances

There is no widely accepted definition of expressivity in musical performance (Fabian, Timmers, & Schubert, 2014, p. xxiii; Juslin & Timmers, 2010, p. 454). In the context of this research, the most suitable definition is that of Leech-Wilkinson (2009), who argues that performers create associations with elements of everyday life by means of musical inflections that bring meaning to the music. He explains:

We can think of these inflections as expressive gestures, expressive because they represent meaning, and gestures because they shape notes over time in the same way, and for the same purpose, that humans use their hands and face to communicate information about the dynamic shape of a process and about its effect. (Leech-Wilkinson, 2009, Chapter 8, ¶13)

Drawing on this notion, he defines expressivity as a change in relation to expectation:

An expressive gesture can be defined as *an irregularity in one or more of the principal acoustic dimensions (pitch, amplitude, duration), introduced in order to give emphasis to a note or chord—usually the start of a note or chord*. Expressive gestures involve sounding notes for longer or shorter, or louder or softer, or in some other way different compared to the local average. Why the local average? Carl Seashore, and many who have followed him, described these irregularities as deviations [...] The problem with that word is not only that it implies deviance, but also that it seems to suppose that there is a proper length, loudness, or pitch for a note. In terms of the score there may be, but as we've repeatedly seen the score is not the music, and nor is a straight performance of it. So 'deviation' from the score is normal, in fact definitive of a musical performance, and it's not the fact that notes are not strictly as notated that generates expressivity. Rather it's how much they differ from their surroundings and from what we've come to accept over the last few moments of listening is *[sic]* the (local) norm. Difference from the score is not what's expressive; change is. (Leech-Wilkinson, 2009, Chapter 8, ¶15, all italics in original)

³ This article does not report analyses of vowel spectra or vowel timbre.

Leech-Wilkinson (2009) also provides some examples of the use of diction as an expressive gesture: he analyses recordings of Schubert solo songs and observes that some singers occasionally emphasize or lengthen consonants (Chapter 8, ¶184-91). In Kathleen Battle's recording of the song *Die Männer sind méchant*, he describes how she elongates the consonant *s* of the word *ist* and emphasizes the consonants *t* from *ist* and *pr* of *Springinsfelt* in the phrase *Er ist ein Springinsfelt* (*ibid.*, ¶185). As he observes:

These are all expressive gestures, evoking sounds from life. Some are onomatopoeic, the 'issst' which we've learned to associate with hate and in particular, a threatening hate that could lead at any moment to violence. Then there are the sudden explosive consonants [...] evoking the sound of sudden violence. (Leech-Wilkinson, 2009, Chapter 8, ¶186)

Leech-Wilkinson also analyses Meta Seinemeyer's recording of the song *Die junge Nonne*. He explains that one of the expressive gestures the singer uses to contrast the first three phrases (about the storm) with the following one (comparing the darkness of the night with death) is an incisive articulation of consonants: the keywords *Wipfel* (treetop), *Balken* (beams), and *Donner* (thunder) "are hit hard through initial consonants sung at full amplitude [...] The speech analogy is obvious: spitting-out sounds evoke anger in speech and by analogy the fury of the storm" (Leech-Wilkinson, 2009, Chapter 8, ¶190).

Sundberg (2000) and Sundberg et al. (1994) also point out that the way one pronounces the sung text can contribute to expressivity. Excerpts from opera and Lieder were recorded by a professional singer specifically for the latter study, sung "in two deliberately contrasting fashions, (1) as in a concert situation and (2) in an emotionally neutral way" (Sundberg et al., 1994, p. 81). After a listening test, in which experts evaluated the recordings, some excerpts were analysed acoustically. With reference to variations of sound level between the neutral and the expressive versions, the authors observe that "[t]he source of a great variability was mostly an emphatic pronunciation of consonants causing a great amplitude modulation in the expressive versions" (*ibid.*, p. 86).

In a subsequent article, Sundberg (2000) analyses the emotional expressivity of the excerpts recorded for the Sundberg et al. study (1994). Among other elements, he analyses the duration of tones, which he defines as the duration from the vowel onset in a syllable to the vowel onset in the next syllable. This definition of tone duration is also used as the definition of syllable. Sundberg (2000, pp. 105-107) observes that the singer emphasized important words by means of lengthening the stressed syllable and the unstressed syllable preceding the stressed one. The lengthening of the unstressed syllable preceding the stressed one is exemplified by the last two words of the phrase *Und beschworst darin die Bösen* (And you banished the evil [spirits], from *Zueignung*, Op. 10 No. 1, by Richard Strauss):

Here, the word 'Bösen' (evil) was perceived as emphasized, which seems logical from a semantic point of view. Although appearing in an unstressed upbeat position in the bar, the syllable (d)'ie B'(ösen) was clearly lengthened, while the syllable (B)'ös'(en) was slightly shortened in the expressive version. (Sundberg, 2000, p. 106)

Here one notices the application of Sundberg's definition of syllable: the "unstressed syllable" does not refer to the *die* that precedes *Bösen*, but to the vowel *ie* plus the initial *b* of *Bösen*. The "stressed syllable" refers to *ös*, and not to *Bö*, which is the stressed syllable of *Bösen*. Sundberg (2000, p. 105) shows the duration ratio of each phoneme in a figure and, with reference to the "unstressed syllable", explains: "It can be seen that the lengthening concerned the consonant [b] rather than the vowel preceding it" (Sundberg, 2000, p. 106). In the final discussion, he adds: "If sung syllables were defined as in orthography, such lengthened consonants would belong to the stressed syllable, so the lengthening would occur

on the stressed syllable. However [...] this definition of syllables does not apply to singing” (Sundberg, 2000, p. 111).

Later I will discuss these studies by Leech-Wilkinson (2009), Sundberg (2000), and Sundberg et al. (1994), in a comparison with the results of the present study.

Diction and expressivity in writings on choral conducting

As mentioned earlier, Western musical notation is not precise in relation to the sung text: because it represents the duration of a syllable as a whole, notation ignores the many sounds that may constitute a syllable. In this way, notation omits not only the exact instant of articulation of the vowels of a diphthong or of the initial and/or final consonants of a syllable, but also the duration, the dynamics, and the timbre of each of these sounds. Analysis of writings on choral conducting show that some authors offer suggestions and/or examples concerning these diction-related choices. The principal results of the analysis are as follows.

In the letters addressed to the choirs he conducted, Robert Shaw describes some of his choices regarding the different sounds of the text. He explains, for example, that initial consonants must be anticipated, that is, articulated “*ahead* of the beat (or beat division) assigned to the syllable” (Blocker, 2004, p. 102), so that the vowel will sound exactly at the beginning of the beat. Final consonants and the unstressed vowel of diphthongs will sound at the last beat (or fraction of beat) assigned to the syllable, yet they should not occupy an accented part of it. Nasal consonants and unstressed vowels must have their dynamics “appreciably increased in order to be heard at all” (*ibid.*, p. 102).

Shaw also writes about “speech as color” (Blocker, 2004, p. 105), drawing attention to the expressive potential of the sounds of the text in view of the variety of timbres of different vowels and consonants. He states that “[t]he color combinations available to the human voice through language literally are infinite” (*ibid.*, p. 106), and concludes:

The choral art of our time⁴ has not even begun to understand and utilize text and enunciation as the consummate conjurers of musical color and timbre. Our attention has been centered upon either the ‘rules’ or the ‘message’—and too frequently these have led to the suffocation of the spirit. We have forgotten that dictionaries offer but silhouetted or graven images of language’s meaning. Poetry never quite survives analysis; dissection ends in inquest; and definitions may be most comforting to the deaf. (Blocker, 2004, p. 106)

Kurt Thomas discusses German diction for choral singing and details some procedures that are specific to this language. For instance, he stresses that, after a short vowel, consonants must be anticipated, thus avoiding the elongation of the vowel, which can distort the meaning of a word. In the word *Sonne* (sun), for example, the *n* must sound immediately after the beginning of the last accented beat (or beat division) assigned to the first syllable; when this syllable has only one beat, the consonant is articulated instantly after the vowel onset (Thomas, 1979, pp. 90-91). He goes on to explain that this is also valid in the case of final short and unstressed vowels, for instance in words such as *gute* (good) or *Ruhe* (rest): when singing *Ruhe wohl* (rest peacefully), for example, the *w* of *wohl* is anticipated, just like the *n* in the word *Sonne* (*ibid.*, pp. 95-96).

Wilhelm Ehmann (1981, pp. 44-45) highlights the importance of careful diction as a means of revealing the sense of the text, giving examples of how the sound of certain words can illustrate their meanings. Such sonic expressivity of words manifests its potential when the

⁴ This letter by Robert Shaw was written in December 1979 (Blocker, 2004, p. 100).

pronunciation is connected with mental images about the meaning of the word — that is to say, when the singer is involved both with the sound and the meaning of the word. Nevertheless, Ehmann stresses that this kind of pronunciation does not aim to dramatize the language nor to add elements that are strange to it, but to extract the maximum from the sonority of the language itself, especially in the case of German.

Simon Halsey has a different approach, by and large: he tends to highlight the decision-making process involved in the performance. As regards diction, he describes his choices concerning the timbre of vowels, the instant of articulation of final consonants or of unstressed vowels of diphthongs, the duration and dynamics of consonants, among others (Halsey, 2011, pp. 88-89, 103). He suggests that in strophic songs the meaning of each verse can be illustrated not only by means of subtle variations in tempo, dynamics, phrasing, or timbre, but also through diction-related choices concerning, for example, the duration and dynamics of vowels and consonants (*ibid.*, pp. 97-104). He also writes about the expressive potential of consonants and exemplifies his choices regarding their duration and dynamics depending on the expressive effect he aims to achieve (*ibid.*, pp. 210-211).

Wilhelm Ehmann and Frauke Haasemann also suggest that consonants can be used to contribute to expressivity (1990, p. 70). They argue that consonants give rhythm to the German language and that consonant clusters create textual pictures, thus enhancing the expressive potential of the language. They describe how consonants can be emphasized or lengthened to illustrate or reinforce the meaning and the expression of the text. For instance, about the first movement of Brahms' *Ein deutsches Requiem*, they write:

In this composition — which has the indication 'mit Ausdruck' and whose vocal parts often bring the request 'expressivo' — an important task falls to the consonants, since their careful treatment 'paints' and lightens the pictures of the text. [...] One can reinforce, for example, the consonants *tr* or *fr* with a diaphragm impulse, thus originating an accumulation prior to the next vowel and with it the desired increase in expression, as a sign of greater inner involvement: *tr'agen, getr'östet, Tr'änen, Fr'eude*.⁵ (Ehmann & Haasemann, 1990, p. 159)

Some of the conductors I interviewed provided examples of situations in which diction can be used expressively. Georg Christoph Biller suggests: "For instance, in the word *brummen* [grumble] one can stress the *m*, in order to make really clear this emotion".⁶ Peter Neumann explains that one can "reinforce the key consonant of a word, when the expression demands".⁷ Stephen Cleobury suggests that "if it is very expressive: *the mother* [sings, slightly lengthening the *m*], you can actually use the length of that consonant as an expressive device". John Butt takes the expression *aus Liebe* (out of love) as an example and explains: "You might spend longer on the *L*, as if you were settling into the meaning and emotion of the word".

⁵ "Den Konsonanten fällt in dieser 'mit Ausdruck' überschriebenen Komposition, deren Stimmarte außerdem häufig die Aufforderung 'expressivo' bringen, eine wichtige Aufgabe zu, da ihre sorgfältige Behandlung die Bilder des Textes 'malen' und aufleuchten lassen. [...] Verstärkt man z. B. die Konsonanten *tr* oder *fr* mit einem Zwerchfell-Impuls, so entsteht eine Stauung vor dem folgenden Vokal und damit die gewünschte Steigerung des Ausdrucks, als Zeichen erhöhten inneren Beteiligtseins: *tr'agen, getr'östet, Tr'änen, Fr'eude*" (Ehmann & Haasemann, 1990, p. 159). All translations are by the author.

⁶ "Zum Beispiel, bei dem Wort 'brummen' kann man das 'm' betonen, um dann wirklich diesen Affekt zu verdeutlichen."

⁷ "den Schlüssel-Konsonanten eines Wortes, wenn es der Ausdruck erfordert, verstärken."

Diction and expressivity in recordings of choral works sung in German

Aiming to identify possible uses of diction as an expressive device, I analysed the treatment of diction in recordings of works by Johann Sebastian Bach and Franz Schubert. I analysed six recordings of the movements *Unter deinem Schirmen* and *Trotz dem alten Drachen* from the Bach's motet *Jesu, meine Freude* BWV 227, conducted by Kurt Thomas, Wilhelm Ehmann, Helmuth Rilling, Eric Ericson, Philippe Herreweghe, and John Eliot Gardiner, and four recordings of Schubert's part-song *An die Sonne* D439, conducted by Frieder Bernius, John Eliot Gardiner, Peter Neumann, and Jörg Straube.

Analyses of *Unter deinem Schirmen* were made primarily by listening critically. Analyses of *Trotz dem alten Drachen* and *An die Sonne* were made using Sonic Visualiser; the durations of vowels and consonants were measured via spectrogram visualisations and listening critically aided by the plugins Note Onset Detector and Onset Detection Function (Duxbury, Bello, Davies, & Sandler, 2003; Pertusa & Iñesta, 2009). Durations of consonants were measured from the beginning of phonation — in the case of fricatives, the beginning of friction; in the case of plosives, the beginning of the plosion⁸ — until the moment the consonant is no longer audible. Durations of vowels were measured from the end of the sound of the consonant, rather than the complete vowel formation, to the beginning of the next consonant. Dynamic analyses (i.e. analysis of loudness of vowels or consonants) were based on the information provided by the plugin Loudness (Bullock, 2007).

Unter deinem Schirmen BWV 227/3

Unter deinem Schirmen, third movement of the motet *Jesu, meine Freude* BWV 227, is a five-part ornamented chorale (SSATB), based on a melody by Johann Crüger. The text, transcribed and translated in Table 1, is the second verse of a poem by Johann Franck. It concerns the protection of Jesus, the anger of Satan, the dangers of enemies and of sin, and states, at the end, trust in Jesus.

Table 1. Text and translation of *Unter deinem Schirmen*

Line	Text	Translation
1	<i>Unter deinem Schirmen</i>	Under your shield
2	<i>bin ich vor den Stürmen</i>	I am free from the assaults
3	<i>aller Feinde frei.</i>	of all enemies.
4	<i>Laß den Satan wittern,</i>	Let Satan scent,
5	<i>laß den Feind erbittern,</i>	let the enemy embitter,
6	<i>mir steht Jesus bei.</i>	Jesus stands by me.
7	<i>Ob es itzt gleich kracht und blitzt,</i>	Even if now there be thunder and lightning,
8	<i>ob gleich Sünd und Hölle schrecken:</i>	Even if sin and Hell frighten:
9	<i>Jesus will mich decken.</i>	Jesus will guard me.

The piece is in three sections, A¹A²B, each one comprising three musical phrases; each phrase corresponds to a line of the text. A¹ and A² set different and somewhat contrasting texts to

⁸ When possible, the duration of a plosive was measured from its occlusion, for instance, [g] in the word *Gras* in *An die Sonne*, the occlusion of which marks the end of the preceding vowel. Nonetheless, it is not possible to identify the occlusion when a voiceless plosive is preceded by a silence, or when [d] is preceded by [n], since the occlusion is mixed with the phonation of [n].

the same music: in A¹ lines 1 to 3 refer to Jesus' protection, but when the same music is repeated in A², lines 4 and 5 refer to Satan and the enemy. Analysis of recordings of this piece aimed to identify possible differences in the treatment of diction in A¹ and A². The durations of vowels and consonants were analysed by means of critical listening. Data regarding the overall dynamics of the piece were based on the information provided by the Loudness plugin.

The analysis showed that the first two phrases in A², corresponding to lines 4 and 5 of the text, were sung louder in all six recordings than their counterparts in A¹. In Ehmann's recording the text seems to be articulated in the same manner in A¹ and A². In Thomas' recording, the consonants in A² of [v] of *wittern* and [b] of *erbittern* seem to be sung a little bit louder than the consonants in A¹. In Ericson's and Herreweghe's recordings, the consonants [l] and [s] of *lass* — as well as [v] of *wittern* in the latter — are clearly articulated. In Rilling's recording, the text is articulated energetically in A²: in the words *lass den* the vowels are short and the consonants are long, and the [v] of *wittern* and [b] of *erbittern* are vigorous. Yet Gardiner's recording is the one that most explores diction as a means of differentiating A² from A¹: one can hear the incisive articulation of consonants such as the [l] and [s] of *lass*, [z] of *Satan*, [v] and [t^h]⁹ of *wittern*, and [b] and [t^h] of *erbittern*. Furthermore, basses sing the vowel [a] of *lass* with an open and harsh timbre (which is reinforced by the bassoon) and with vibrato, in such a way that the sonority of lines 4 and 5 (in A²) strongly contrasts with that of lines 1 and 2 (in A¹).

In the B section (lines 7 to 9), Gardiner's recording continues to explore the expressive potential of diction. Critical listening shows that, in the first musical phrase, a crisp and clear articulation of consonants illustrates the meaning of the line *ob es itzt gleich kracht und blitzt*. In the following phrases, this kind of illustration of the text is also achieved by varying the vocal timbre: in the first syllable of the word *schrecken*, the timbre of the vowel [ɛ] is quite open in comparison with the round and soft [ɛ] of *decken*.

Analyses of these six recordings reveal different interpretative strategies, especially in sections A¹ and A². In Rilling's and Gardiner's recordings diction, as well as dynamic variation, is used to illustrate different meanings of the text. This relates to Halsey's suggestions concerning strophic songs, quoted earlier.

The effect of these choices seems, to me, to be related to different perspectives on the tension expressed in the text between Jesus and Satan, trust and fear. The recordings in which A¹ and A² are differentiated principally by varying the dynamics¹⁰ tend to emphasize the strength of Jesus' protection. Withstanding Satan does not lead to disquietude; sin and hell do not really terrify. Trust in Jesus allows one to talk about them with a certain detachment, as if one were far from mundane fears.

By contrast, Rilling's and especially Gardiner's recordings, which use the expressivity of the sounds of the text along with variations in the dynamics, seem to emphasize the conflict between trust in Jesus and fear. In Gardiner's recording, withstanding Satan provokes tension; thunder and lightning do intimidate. Moreover, the terror of hell and sin, as illustrated in line 8 of this recording, sounds real and close, so that we perceive the speaker as a human being with strengths and weaknesses. This disquietude contrasts with the serenity with which faith in Jesus is stated in both the first and the last lines of this movement.

⁹ Since German voiceless plosives are generally aspirated (Boor, Moser, & Winkler, 1969, pp. 22, 104), I decided to use the diacritic [ʰ], which indicates the aspiration, in the phonetic transcriptions.

¹⁰ It is possible that there is some tempo variation between sections A¹ and A², although I did not analyse tempo variations in these six recordings.

Trotz dem alten Drachen BWV 227/5

Trotz dem alten Drachen, the fifth movement of *Jesu, meine Freude* BWV 227, is a five-part (SSATB) free-setting of the same melody by Johann Crüger used in other movements of this motet.¹¹ The text, transcribed and translated in Table 2, is the third verse of Johann Franck's poem. As Hofmann (2006, p. 128) explains, here "[t]he forces hinted at in the second chorale strophe [that is, in the third movement of the motet] are unchained"¹²: the one who trusts in Jesus confronts the peril of death, the fear of it, and the world's rage, defying all this and stating his or her faith and serenity.

Table 2. Text and translation of *Trotz dem alten Drachen*

Line	Text	Translation
1	<i>Trotz dem alten Drachen,</i>	Defy the old dragon,
2	<i>trotz des Todes Rachen,</i>	defy the peril of death,
3	<i>trotz der Furcht darzu!</i>	defy the fear of it!
4	<i>Tobe, Welt, und springe,</i>	Rage, world, and storm,
5	<i>ich steh hier und singe</i>	I stay here and sing
6	<i>in gar sichrer Ruh.</i>	in safe tranquility.
7	<i>Gottes Macht hält mich in Acht;</i>	God's power keeps me secure;
8	<i>Erd und Abgrund muss verstummen,</i>	Earth and abyss must be silent,
9	<i>ob sie noch so brummen.</i>	even if they still grumble.

The first three lines of this verse begin with the word *trotz*, which can refer to the preposition *trotz*, meaning 'despite', the noun *Trotz*, meaning 'defiance' or 'obstinacy', or the imperative of the verb *trotzen*, which can be translated as 'defy', 'oppose', or 'be obstinate'. In Franck's text, *trotz* is probably used as a preposition. Nevertheless, Bach's setting of the word — especially when it is followed by a rest — illustrates, rather, an imperative. I have translated it thus as it seems to reflect better the conviction expressed by the music.

Hofmann writes that this movement represents "a kind of scene of impetus and fury that hardly finds an equivalent in the motet literature"¹³ (2006, p. 128). In performance, this fury can be illustrated not only through variations in dynamics or tempo, but also through expressive diction. The following analyses of recordings show that, occasionally, consonants are lengthened, thus reinforcing the vigorous character of the piece.

At the beginning of this movement, the word *trotz* is sung twice with interposed rests, thus establishing the impetuous sense of Bach's composition. With the software Sonic Visualiser, I measured the duration of the consonant cluster [t^hr]¹⁴ of *trotz* in the first two bars of the piece (bars 147 and 148).¹⁵ Table 3 presents the mean beats per minute (bpm) in this passage,¹⁶ the durations in milliseconds of [t^hr], and the percentage of the duration of

¹¹ It is a free setting because Crüger's original melody is hardly recognisable in this movement.

¹² "Die Kräfte, die sich in der zweiten Choralstrophe andeuten, werden entfesselt" (Hofmann, 2006, p. 128).

¹³ "eine Art Sturm- und Furienszene, die in der Motettenliteratur wohl nicht leicht ihresgleichen findet" (Hofmann, 2006, p. 128)

¹⁴ According to the *Siebs Deutsche Aussprache* (Boor, Moser, & Winkler, 1969, p. 104), German voiceless plosives are aspirated also when followed by [r] or [l].

¹⁵ Bar numbers refer to motet BWV 227 as a whole.

¹⁶ When comparing durations of vowels and consonants, the measure in milliseconds — to some extent 'absolute' — must be relativized, since different interpretations of a work have rarely identical tempi. Therefore,

the mean bpm that is occupied by [t^hr].¹⁷

Table 3. Duration of [t^hr] of *trotz* in bars 147 and 148 of *Trotz dem alten Drachen*

Conductor	Mean bpm	[t ^h r] bar 147	Percentage of mean bpm	[t ^h r] bar 148	Percentage of mean bpm
Thomas	80	158	21.1%	176	23.5%
Ericson	88	259	38%	245	35.9%
Ehmann	93	242	37.5%	277	42.9%
Rilling	94	272	42.6%	233	36.5%
Herreweghe	99	148	24.4%	226	37.3%
Gardiner	101	307	51.7%	273	46%

The consonant cluster [t^hr] (bar 147) has the shortest durations in Thomas' and Herreweghe's recordings, and the longest in Gardiner's. Although Gardiner and Herreweghe use similar tempi, the [t^hr] is twice as long in Gardiner's recording as in Herreweghe's. Conversely, the [t^hr] has almost the same duration in Thomas' and Herreweghe's recordings, even though Thomas uses the slowest and Herreweghe the second fastest tempo. By comparison, both the durations of [t^hr] and the tempi used are intermediate in Ericson's, Ehmann's, and Rilling's recordings. In the repetition of the word *trotz* (bar 148), the longest durations of [t^hr] can be heard in Ehmann's and Gardiner's recordings, and the shortest in Thomas'. All the recordings reveal a (generally almost imperceptible) variation in the duration of [t^hr] in bar 148, as compared to that in bar 147; the biggest difference is found in Herreweghe's recording, in which the [t^hr] in bar 148 lasts roughly 50% more than in bar 147.

When listening critically to this passage, I notice that in Gardiner's recording the word *trotz* is emphasized in both bars 147 and 148. In Herreweghe's recording, the word *trotz* is emphasized in bar 148. It does not stand out from its context in the other recordings.

The next passages analysed refer to the word *Drachen*, sung in bars 149-150 and 151. Table 4 shows the mean beats per minute (bpm), the durations in milliseconds of the consonants [dr] and [x] (*ch*) of *Drachen* in bars 149-150, and the percentage of the duration of the mean bpm that is occupied by these consonants.

in each analysed passage of a recording, I calculated the mean beats per minute (bpm); in the tables, recordings are listed according to this mean. Bpm were calculated with Sonic Visualiser. The mean bpm shown in Tables 3, 4 and 5 refers to that used from the beginning of bar 147 to the last crotchet of bar 154.

¹⁷ These percentages are approximate. This happens not only because they were calculated in relation to the mean bpm — and not the exact bpm of the first crotchet of bars 147 and 148 —, but also because most of the duration of [t^hr] sounds before the beat (the beginning of the beat was identified according to the instrumental attack). This also refers to Tables 4, 5, 6, 8, and 9.

Table 4. Duration of [dr] and [x] of *Drachen* in bars 149-150 of *Trotz dem alten Drachen*

Conductor	Mean bpm	[dr] bar 149	Percentage of mean bpm	[x] bar 150	Percentage of mean bpm
Thomas	80	137	18.3%	162	21.6%
Ericson	88	224	32.8%	206	30.2%
Ehmann	93	171	26.5%	152	23.6%
Rilling	94	141	22.1%	134	21%
Herreweghe	99	176	29%	173	28.5%
Gardiner	101	216	36.4%	317	53.4%

The consonant cluster [dr] (bar 149) has the shortest durations in Thomas' and Rilling's recordings, intermediate durations in Ehmann's and Herreweghe's, and the longest durations in Ericson's and Gardiner's recordings. In each of these pairs of recordings, the tempi are dissimilar but the differences in the durations of [dr] are small. For instance, Ericson and Gardiner use the second slowest and the fastest tempi respectively, but in the two recordings the [dr] has almost the same duration.

The long duration of the consonant [x] (bar 150) in Gardiner's recording calls attention to itself; notwithstanding Gardiner's fast tempo, the [x] lasts more than twice as long as the mean duration of the four shortest [x] consonants in Rilling's, Ehmann's, Thomas', and Herreweghe's recordings. Furthermore, the friction noise of this consonant can be clearly heard in Gardiner's recording, in bar 150 as well as 151.

I analysed the durations of the consonants [dr] and [x], and also of the vowel [a] of *Drachen* in bar 151. Here, the word is sung with notes of smaller time values than in the previous passage: each syllable lasts for a quaver rather than a minim for *Dra* and a crotchet for *chen* (bars 149-150). Yet the consonant cluster [dr] is not necessarily shorter in bar 151, as it is in Ericson's, Ehmann's, and Herreweghe's recordings; it is longer in those of Thomas, Rilling, and Gardiner. By comparison, the consonant [x] is shorter in bar 151 than in 150 in all recordings, as shown in Table 5.

Table 5. Duration of [dr], [a] and [x] of *Drachen* in bar 151 of *Trotz dem alten Drachen*

Conductor	Mean bpm	[dr] bar 151	Percentage of mean bpm	[a] bar 151	Percentage of mean bpm	[x] bar 151	Percentage of mean bpm
Thomas	80	165	22%	199	26.5%	136	18.1%
Ericson	88	182	26.7%	176	25.8%	91	13.3%
Ehmann	93	122	18.9%	156	24.2%	128	19.8%
Rilling	94	164	25.7%	161	25.2%	69	10.8%
Herreweghe	99	144	23.8%	165	27.2%	119	19.6%
Gardiner	101	224	37.7%	107	18%	175	29.5%

Comparison of the durations of [x] and [dr] (bar 151) shows that [x] is in general shorter than [dr], except in Ehmann's recording, in which [x] is just a little bit longer than [dr]. The vowel [ɑ] lasts longer than the consonants [dr] and [x] in Thomas', Ehmann's, and Herreweghe's recordings. In Ericson's and Rilling's, it has virtually the same duration as [dr], but approximately twice the duration of [x]. Yet again Gardiner's recording differs from the others: in addition to having the shortest vowel, it is the only recording in which the [ɑ] is shorter than both consonants, lasting less than half of [dr] and less than two-thirds of [x].

When listening critically to both the passages in which *Drachen* is sung, I hear a stress on the syllable *Dra* in bar 149 of Ericson's recording, while in Gardiner's recording the word *Drachen* is stressed in both bars 149-150 and 151 principally because of the duration of the consonant [x], but also because of the duration of [dr]. This word does not sound stressed in the other recordings.

The durations shown so far suggest that a slower tempo — in which there would be more time for each vowel and consonant to sound — does not necessarily result in longer consonants, nor does a faster tempo prevent their being lengthened. Rather, the duration of consonants seems to be related to the conductor's choices regarding the text. In Gardiner's recording, for example, it is quite clear that the choice was made to highlight important words by using long and incisive consonants, thus reinforcing the energetic and impetuous character of this movement.

The final passage to be analysed concerns the duration and dynamics of vowels and consonants in the words *ich* and *steh*, from the fifth line of the text (bars 167-168). Table 6 shows the durations in milliseconds of [ɪ] and [ç] (*ch*) of *ich*, and of the consonant cluster [ʃtʰ] (*st*) of *steh*. The recordings have been reordered in this table because, in all of them, there is a slowing of the tempo used at the beginning of the movement.¹⁸

Table 6. Duration of [ɪ] and [ç] of *ich* in bar 167 and of [ʃtʰ] of *steh* in bar 168 of *Trotz dem alten Drachen*

Conductor	Mean bpm	[ɪ] bar 167	Percentage of mean bpm	[ç] bar 167	Percentage of mean bpm	[ʃtʰ] bar 168	Percentage of mean bpm
Thomas	76	676	85.7%	231	29.3%	286	36.2%
Ehmann	83	417	57.7%	207	28.6%	222	30.7%
Ericson	84	581	81.4%	155	21.7%	333	46.6%
Rilling	88	634	93%	238	34.9%	252	37%
Herreweghe	90	463	69.4%	220	33%	396	59.4%
Gardiner	95	452	71.5%	244	38.6%	372	58.9%

The longest vowels are to be heard in the pronoun *ich* in Thomas' and Rilling's recordings, while the consonant [ç] lasts approximately a third of the vowel. Although Ehmann and Ericson use almost identical tempi, Ehmann's recording has the shortest [ɪ], with the [ç] lasting half of the vowel duration, while Ericson's has the third longest vowel and the shortest

¹⁸ The mean bpm shown in Table 6 refers to that used from the second crotchet of bar 166 to the last crotchet of bar 169.

consonant at just over a quarter of the [ɪ]. In Herreweghe's recording, [ç] lasts just under half and in Gardiner's recording more than half of the length of the vowel. Data provided by the plugin Loudness show that in Ehmann's and Gardiner's recordings, despite having the shortest vowels — and, to a lesser degree, in Herreweghe's too — the onset of the [ɪ] vowel is *forte* (even though in Ehmann's recording it is not as loud as the preceding bar). By comparison, its onset is soft in Thomas', Ericson's, and Rilling's recordings. It stays soft over its full duration in Ericson's recording; it has a *crescendo* in Rilling's; and only in Thomas' recording does the *crescendo* reach the *forte* of the preceding bar.

The consonant cluster [ʃtʰ] of *steh* (bar 168) has the longest durations in Herreweghe's and Gardiner's recordings, corresponding to more than a quaver, at three-quarters of the duration of the [ɪ] in the previous bar. Its dynamic is, by and large, related to that of the passage: in the recordings in which *ich steh* is sung *forte* throughout, the consonants are *forte*; in those in which it is sung *piano* with a *crescendo*, [ʃtʰ] is sung with a *crescendo*.

The dynamics of the [e:] of *steh* (bar 168) were compared, using the plugin Loudness, with the [ɪ] of *ich* (bar 167) and with the general dynamic contour of the previous line *tobe, Welt, und springe* (bars 166-167). The comparison revealed a *crescendo* on [e:] in relation to [ɪ] in Ericson's recording, although it does not reach the *forte* of the preceding line. In Thomas', Rilling's, and Ehmann's recordings, the [e:] sounds louder than the [ɪ] and does not take long to nearly reach the *forte* of the previous line. There is a slow *crescendo* in the [e:] in Herreweghe's recording, which reaches a dynamic louder than that of the previous line. In Gardiner's recording, the [e:] is attacked immediately at a dynamic a bit louder than that used in the preceding two bars.

Ericson and Gardiner take the most dissimilar approaches to this passage. In Ericson's recording, *ich steh hier und singe* (line 5, bars 167-169) is sung more softly, despite its *crescendo*, than the surrounding bars. The *ich* in particular — sung *piano* with a long vowel — creates a contrast to the previous passage and anticipates the calm atmosphere of the next line, *in gar sichrer Ruh*. In Gardiner's recording, with its long consonants and *forte* dynamics, the assertion of line 5 sounds categorical. Rather than contrasting with the previous line, it persists with a challenge: "The world can rage, for I will definitely stay here and sing." The contrast comes in the next line, *in gar sichrer Ruh*. When listening critically to this passage, I hear a clear breath — a short pause that sounds emphatic and, at the same time, almost as though it were a sigh of relief — between the first and the second beats of bar 175, just before the repetitions of *in gar sichrer Ruh*, and then a sonorous [z] that highlights the word *sicher* at the beginning of bar 176.

The effect of these choices seems, as mentioned earlier, to point to different perspectives in relation to the theme presented in this and also the third movement of this motet: trust in Jesus and fear. In Ericson's recording, the lines *ich steh hier und singe in gar sichrer Ruh* sound tranquil, in spite of the unison motive sung louder in bar 170; one could imagine someone who sings as though praying, in a safe and protected situation, far from the dangers of the world. In Gardiner's recording, these same lines show someone who is resisting attack — depicted from the beginning of the movement, for instance, by means of the long consonants of *Drachen* — while it is happening. To be sure, the line *in gar sichrer Ruh* does sound tranquil, but this tranquillity sounds as though it is experienced by someone confident of winning the battle rather than far away from it.

An die Sonne D439

The part-song *An die Sonne D439* by Schubert was written for a quartet of mixed voices (SATB)

with piano accompaniment. Schubert used as his text the first three verses of a 12-verse poem by Johann Peter Uz. These three verses, transcribed and translated in Table 7, concern the sun, nature, and death.

Table 7. Text and translation of *An die Sonne*

Text	Translation
<i>An die Sonne</i>	To the sun
<i>O Sonne, Königin der Welt, die unser dunkles Rund erhellt, in lichter Majestät; erhab'nes Wunder einer Hand, die jene Himmel ausgespannt und Sterne hingesät!</i>	O Sun, queen of the world, who illuminates our dark world, in brilliant majesty; sublime miracle from a hand, that spread out the heavens and sowed the stars!
<i>Noch heute seh' ich deinen Glanz, mir lacht in ihrem Blumenkranz noch heute die Natur. Der Vögel buntgefiedert Heer singt morgen mir vielleicht nicht mehr im Wald und auf der Flur.</i>	Today I still see your brightness, with its garland of flowers today nature still smiles at me. The host of birds with colourful feather may tomorrow not sing to me again in the woods and in the meadows.
<i>Ich fühle, daß ich sterblich bin, mein Leben welkt wie Gras dahin, wie ein verschmachtet Laub. Wer weiß, wie unerwartet bald des Höchsten Wort an mich erschallt: Komm wieder in den Staub!</i>	I feel that I am mortal, my life withers like grass, like languishing leaves. Who knows how unexpected and soon the word of the Almighty will sound to me: Return to dust!

I analysed the treatment of diction in the first three lines of the poem's third verse in four recordings of this part-song. They are sung for the first time in bars 60-69, presenting the recognition that life is finite in a tranquil manner: the notes are longer than in previous passages, and major harmonies, *pianissimo* dynamics, and voices in homophony predominate, providing a serene atmosphere. The same lines are repeated in bars 78-90 but this time the serene atmosphere is abandoned at the end of the first line, and the second and third lines acquire a dramatic character. The line *mein Leben welkt wie Gras dahin* is no longer homophonic and has greater harmonic tension — at the beginning of bar 84, for instance, the word *Gras* is emphasized by means of a minor ninth chord, a *crescendo*, and a note of greater time value. In the next line, harmonic tension is produced when the word *verschmachtet* ("languishing": bars 88-89 including upbeat) is sung. This word sounds dramatic not only because it lasts three times as long as in bar 68 and its stressed syllable has a *crescendo*, but also, more importantly, because of the dissonances created by the sequence of suspensions in the soprano line. The line ends with a *diminuendo*, the weakening it expresses corresponding to the meaning of *verschmachtet*.

Considering the different contexts in which the lines *mein Leben welkt wie Gras dahin*, *wie ein verschmachtet Laub* (bars 65-69 and 83-90 including upbeat) are sung, analyses of the recordings aimed to identify differences in the diction used in two important words: *Gras* (bars 66 and 84) and *verschmachtet* (bars 68 and 88-89 including upbeat). Analysis shows

that these words are articulated differently when they are sung for the first and second times: in virtually all the recordings, the consonant clusters [gr] of *Gras* and [ʃm] (*schm*) of *verschmachtend* are longer when repeated in bars 84 and 88-89.

Table 8 shows the mean bpm of the two passages in which the word *Gras* is sung,¹⁹ the durations in milliseconds of [gr], and the percentage of the duration of the mean bpm that is occupied by [gr].

Table 8. Duration of [gr] of *Gras* in bars 66 and 84 of *An die Sonne*

Conductor	Mean bpm	[gr] bar 66	Percentage of mean bpm	Mean bpm	[gr] bar 84	Percentage of mean bpm
Straube	88	100	14.7%	79	197	26%
Neumann	92	216	33.1%	78	325	42.3%
Gardiner	94	243	38.1%	87	359	52%
Bernius	108	127	22.8%	84	278	38.9%

Bernius' recording of these passages differs most from the others in that the duration of the initial consonant cluster [gr] of *Gras* in bar 84 is more than twice as long as the [gr] in bar 66. The tempo also slows down more steeply between the two occurrences of the word than in the other recordings, but the longer [gr] in bar 84 is not necessarily a direct result. For instance, Gardiner's recording has the fastest tempo and the longest [gr] in bar 84, Neumann's the slowest tempo and the second longest [gr], while Straube's has virtually the same tempo as Neumann's, but the shortest [gr]. As discussed above, such comparisons suggest that slow tempi and long consonants, and fast tempi and short consonants are not necessarily related, even though consonants need time to be articulated and, above all, lengthened.

In all four recordings, the word *Gras* seems to be emphasized in bar 84, and one contributory element is the lengthening particularly of the [r] in [gr]. Differences between the durations of [gr] in each recording are affected by musical context. When listening critically to compare Neumann's and Straube's recordings of bar 84 in isolation, for example, I have the impression that *Gras* is not emphasized in the latter. Yet, when listening to Straube's complete recording, I notice a lengthening of [gr] and an emphasis on this word that corresponds with his interpretation of the work as a whole.

In the two passages in which the word *verschmachtend* is sung, there is a reduction in tempo, shown in Table 9.²⁰ The recordings have been reordered in this table because Neumann's recording of these passages has the slowest tempo.

¹⁹ Bpm indicated for bar 66 refer to the mean bpm of bars 65 and 66. Bpm indicated for bar 84 refer to the mean bpm from the last crotchet of bar 82 to the third crotchet of bar 84.

²⁰ Bpm indicated for bar 68 refer to the mean bpm from the last crotchet of bar 67 to the third crotchet of bar 68. Bpm indicated for bar 88 refer to the mean bpm from the second crotchet of bar 87 to the second crotchet of bar 89.

Table 9. Duration of [ʃm] (*schm*) of *verschmachtend* in bars 68 and 88 of *An die Sonne*

Conductor	Mean bpm	[ʃm] bar 68	Percentage of mean bpm	Mean bpm	[ʃm] bar 88	Percentage of mean bpm
Neumann	83	336	46.5%	70	476	55.5%
Straube	89	359	53.3%	79	464	61.1%
Gardiner	93	429	66.5%	74	452	55.7%
Bernius	99	394	65%	72	592	71.1%

In three of the four recordings the percentages that the [ʃm] of *verschmachtend* occupies in relation to the mean bpm are greater when this word is sung for the second time. The longest [ʃm] in bar 88 is in Bernius' recording, although Neumann's recording has the greatest increase, from bar 68 to bar 88, in the percentages that [ʃm] occupies in relation to the mean bpm. Meanwhile Gardiner's recording has a decrease in the equivalent percentages, even though the duration in milliseconds of [ʃm] in bar 88 is longer. It is worth noting that in all four recordings there is a *ritenuto* between *ver* and *schma*, from bar 87 to bar 88. Like the lengthening of [ʃm], the tempo reduction relates to the musical setting of the text, with dissonant chords evoking the moment when death is acknowledged. The tempo reduction certainly enables the consonants to be lengthened, especially the [ʃ] which is anticipated (i.e., articulated before the beat) in all the recordings and needs time to sound before the attack of the first beat of bar 88. Nevertheless, as discussed above, tempo reduction is not the only explanation for the lengthening of consonants.

The dynamic outline provided by the plugin Loudness shows changes in the dynamics of the line *wie ein verschmachtend Laub* (bars 87-90), which suggest that the increase in the duration of [ʃm] in bar 88 in relation to bar 68 may not be the decisive factor for this consonant cluster to sound emphasized. For instance, in Straube's recording bars 87-90 are sung *piano* and with almost no *crescendo*, so that although the [ʃm] is long, it is not emphasized. In Gardiner's recording, the percentage of [ʃm] in relation to the mean bpm in bar 88 is smaller than in bar 68. Nevertheless the *schma* of *verschmachtend* sounds emphasized mainly because of the *crescendo* from bar 87 to the first beat of bar 88. Bernius' recording has a long [ʃm] in bar 88 but also a *crescendo* and *decrescendo* in bars 87-88 that emphasize the *schma*. Neumann uses soft singing but accents the *schma* on the first two beats of bar 88.

These situations show how consonants can be emphasized both by means of duration and dynamics: in Straube's, Neumann's, and Gardiner's recordings, the [ʃm] in bar 88 has similar durations and occupies similar percentages of the mean bpm, yet in Neumann's and Gardiner's recordings the [ʃm] seems to be emphasized more through the use of dynamics than duration, while in Straube's recording they are not emphasized.

I did not measure the duration of [xtʰ] (*cht*) in the word *verschmachtend*, because no quantitative comparison could be made: it is not sung homophonically in bar 89 as it is in bar 68. However, critical listening reveals that, by and large, the consonant [x] in bar 89 sounds much softer than in bar 68, being almost inaudible when sung by women's voices in Neumann's and Straube's recordings, which, as described above, use the softest dynamics. Gardiner's recording is an exception: in it, the [x] is clearly heard, in spite of the *diminuendo*.

Analysis of recordings shows that different strategies were used to convey the text *mein Leben welkt wie Gras dahin, wie ein verschmachtend Laub* (bars 83-90 including upbeat).

Bernius, Gardiner, and Neumann use both a lengthening of the consonants [gr] and [ʃm] of *Gras* (bar 84) and *verschmachtend* (bar 88) and a *crescendo* in the stressed syllable, particularly in *verschmachtend*, that create contrast with what has gone before and is to come. These elements impart an almost theatrical representation of the distress experienced when confronting the finiteness of life. The recordings thus reinforce the dramatic character of the passage, which contrasts with the serene atmosphere of the previous line *Ich fühle, daß ich sterblich bin* (bars 79-82). In Straube's recording, which uses soft dynamics almost exclusively, the serenity of that line is maintained for the recognition of death. While *Gras* has longer initial consonants and a slight dynamic increase, thus becoming more expressive, the line does not sound as dramatic as in the other recordings. *Verschmachtend Laub* is sung as though by someone approaching a peaceful death: although the [ʃm] is long it is not stressed and the softness of the [x] represents the weak articulation of someone whose forces are spent. Thus Straube's recording does not reinforce, but to a certain extent reduces the dramatic character of this passage, bringing to it a serene quality.

DISCUSSION

The analyses of recordings described above show situations in which text diction is used as an expressive device, generally by means of emphasizing or lengthening consonants, but occasionally also by producing them more quietly. Such manipulations of the durations and dynamics of consonants occur in passages that have a vigorous or energetic character, such as the movements of Bach's motet, or are dramatic, such as Schubert's part-song.

In general, emphasis and lengthening occur on the initial consonant, or consonant cluster, of the stressed syllable of an important word. Gardiner's recording of Bach's motet also reveals emphases on the initial consonant of the unstressed syllable that follows the stressed one. Neumann's and Straube's recordings of the Schubert's part-song illustrate the use of softened consonants as an expressive device. Diction can also be manipulated to illustrate the character of different texts sung to the same music, as exemplified in Rilling's and Gardiner's recordings of *Unter deinem Schirmen*. The degree to which consonants are emphasised or lengthened varies between recordings, as one would expect, and does not seem to be influenced by tempo, given that consonants are not necessarily shorter when tempi are faster nor longer when tempi are slower.

Some of the writings on choral conducting that I analysed and interviews I conducted with conductors mentioned lengthening or articulating energetically the initial consonant(s) of stressed syllables, and those of subsequent unstressed syllables to emphasize important words. My results are consistent also with Leech-Wilkinson's analyses of Schubert solo songs (Leech-Wilkinson, 2009, Chapter 8, ¶185-90), showing that these strategies are applicable to choral music. My results do not corroborate Sundberg's findings (2000, pp. 105-107), however, particularly in relation to his definition of "syllable" and his conclusion that unstressed syllables preceding stressed syllables are lengthened. For example Thomas (1979, pp. 95-96) insists that final, unstressed vowels must not be lengthened, and indeed critical listening to Gardiner's recording of *Trotz dem alten Drachen* reveals that the initial consonant cluster [dr] of *Drachen* is emphasized but not final *en* of *alten* that precedes it.

The manipulation of text diction observed in some of the recordings analysed can be understood as an expressive gesture, as defined by Leech-Wilkinson (2009, Chapter 8, ¶13-15, quoted earlier): firstly, a lengthened consonant can shape a note and create meaning; secondly, this lengthening distinguishes the consonant from its surroundings, thus resulting in a change, in respect to the expectation created by the performance, that is perceived as

expressive. These expressive gestures can be connected to the speech sounds used to express such meanings (or emotions, or moods), as mentioned by Leech-Wilkinson (2009, Chapter 8, ¶183) and in studies of emotional speech. In some of the more vigorous moments of *Unter deinem Schirmen* and *Trotz dem alten Drachen*, emphasizing or lengthening consonants — which can result in a shortening of the vowel, as in Gardiner's recording — produces an effect similar to that observed by Trojan (1952, pp. 183-184) in which consonants predominate in emotions such as defiance, and Fónagy's descriptions (1991, p. 154) of aggressive attitudes shown by increasing the duration of consonants and reducing those of vowels. Conversely, the lengthening of consonants in the dramatic passage from *An die Sonne* can be compared to the characterisation of sorrow as described by Sundberg (1987, p. 151). Moreover, the soft dynamics of the consonant [x] shows the feeble articulation one would expect of a languishing person.

When considering these expressive gestures in their musical contexts, I notice that the performance choices made by each conductor can reveal a range of potential meanings for each work. Analysis of diction in recordings of *Unter deinem Schirmen*, for example, reveals different ways of interpreting the tension between Jesus and Satan expressed in the text: some recordings seem to reinforce the protection of Jesus, while others tend to emphasize the conflict between trust and fear. This suggests that manipulations of text diction not only reinforce the meaning of a piece, but occasionally also create different meanings. Such results relate to Leech-Wilkinson's observation of "how much the meaning of a composition can be transformed by a performance, through the signification of the expressive gestures that the performer deploys" (Leech-Wilkinson, 2009, Chapter 8, ¶196). Of course other elements not analysed in the present study also contribute to expressivity and the creation of meaning: Juslin (2003, p. 280) states that "expression is a multi-dimensional phenomenon" and Leech-Wilkinson's (2009) analyses of recordings show that performers can use several elements to shape expressivity as desired. Thus manipulations of text diction are certainly not the only, but one tool for shaping expressivity and meaning.

The role that manipulations of text diction can have in the expressivity of choral performance seems to be connected with the intensity and the frequency of use of such manipulations. For instance, both Ericson and Gardiner use diction as an expressive device in *Trotz dem alten Drachen* but in different ways. Ericson manipulates the duration of consonants occasionally, and while it contributes to the expressivity of the performance, it is not so fundamental. By contrast, Gardiner not only uses some quite prolonged consonants, but he also uses them more often. Such manipulations have a central role in expressivity in Gardiner's recording and are as significant — if not more so — than performance elements such as variations in tempo or vibrato that have hitherto received more attention in research on expressivity in performance.

To conclude, the issues discussed above address the ways in which musical performance creates meanings — as Leech-Wilkinson (2009, Chapter 8, ¶196) has written — by means of the expressive gestures performers choose to employ. Thus performance choices can determine the results of a performance and its effects on the listener. Choral conductors should accept the invitation to reflect on what to choose — or, beyond that, to recognise that not to choose is also a choice.

ACKNOWLEDGEMENTS

This paper draws on my PhD research *Dicção, expressividade e escolhas do regente em obras corais em alemão: discutindo relações entre escritos e gravações*, which was conducted at the University of São Paulo (Brazil) and partly at the University of Cambridge (UK). The research was funded by the CAPES Foundation, Ministry of Education of Brazil, Brasília, DF 70.040-020, Brazil, process number 99999.008904/2014-06. I am grateful to the Managing Editor, Prof. Jane Ginsborg, the Action Editor, Dr Philip Fine, and two anonymous reviewers for their helpful comments.

REFERENCES

- Bastian, H.G., & Fischer, W. (2006). *Handbuch der Chorleitung*. Mainz: Schott.
- Blocker, R. (Ed.) (2004). *The Robert Shaw reader*. New Haven: Yale University Press.
- Boor, H. de, Moser, H., & Winkler, C. (Eds.) (1969). *Siebs, deutsche Aussprache: reine und gemäßigte Hochlautung mit Aussprachewörterbuch* (19th ed.). Berlin: Walter de Gruyter.
- Bullock, J. (2007). Libxtract: A lightweight library for audio feature extraction. In *Proceedings of the International Computer Music Conference* (pp. 25-28). Copenhagen: ICMA. Retrieved July 14, 2016, from <http://hdl.handle.net/2027/spo.bbp2372.2007.116>
- Cannam, C., Landone, C., & Sandler, M. (2010). Sonic Visualiser: An open source application for viewing, analysing, and annotating music audio files. In *Proceedings of the ACM Multimedia 2010 International Conference* (pp. 1467-1468). Firenze: MM'10. Retrieved March 25, 2015, from <http://www.sonicvisualiser.org/sv2010.pdf>
- Duxbury, C., Bello, J.P., Davies, M., & Sandler, M. (2003). Complex domain onset detection for musical signals. In *Proceedings of the 6th Conference on Digital Audio Effects (DAFx-03)* (pp. 1-4). London: DAFx. Retrieved July 14, 2016, from <http://www.eecs.qmul.ac.uk/legacy/dafx03/proceedings/pdfs/dafx81.pdf>
- Ehmann, W. (1981). *Die Chorführung: II. Das künstlerische Singen* (6th ed.). Kassel: Bärenreiter.
- Ehmann, W., & Haasemann, F. (1990). *Handbuch der chorischen Stimmbildung* (3rd ed.). Kassel: Bärenreiter.
- Emmons, S., & Chase, C. (2006). *Prescriptions for choral excellence: Tone, text, dynamic leadership*. Oxford, New York: Oxford University Press.
- Fabian, D., Timmers, R., & Schubert, E. (2014). Introduction. In D. Fabian, R. Timmers, & E. Schubert (Eds.), *Expressiveness in music performance: Empirical approaches across styles and cultures* (pp. xxi-xxx). New York: Oxford University Press.
- Fónagy, I. (1991). *La vive voix: Essais de psycho-phonétique*. Paris: Payot.
- Garretson, R.L. (1998). *Conducting choral music* (8th ed.). New Jersey: Prentice Hall.
- Halsey, S. (2011). *Schott Master Class Chorleitung: vom Konzept zum Konzert*. Mainz: Schott Music.
- Hammar, R.A. (1984). *Pragmatic choral procedures*. Metuchen, N.J., London: Scarecrow Press.
- Hofmann, K. (2006). *Johann Sebastian Bach: die Motetten* (2nd ed.). Kassel: Bärenreiter.
- Jansens, S., Bloothoof, G., & Krom, G. de (1997). Perception and acoustics of emotions in singing. In *Proceedings of the 5th European Conference on Speech Communication and Technology*, Vol. IV, pp. 2155-2158. Rhodes: European Speech Communication Association. Retrieved September 12, 2016, from <https://www.isca->

- speech.org/archive/archive_papers/eurospeech_1997/e97_2155.pdf
- Juslin, P.N. (2003). Five facets of musical expression: A psychologist's perspective on music performance. *Psychology of Music*, 31 (3), 273-302. doi: 10.1177/03057356030313003
- Juslin, P.N., & Timmers, R. (2010). Expression and communication of emotion in music performance. In P.N. Juslin & J.A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 453-489). New York: Oxford University Press.
- Kaplan, A. (1985). *Choral conducting*. New York: W. W. Norton.
- Leech-Wilkinson, D. (2009) *The changing sound of music: Approaches to studying recorded musical performance*. London: CHARM. Retrieved August 15, 2016, from <http://www.charm.kcl.ac.uk/studies/chapters/intro.html>
- Pertusa, A., & Iñesta, J.M. (2009). Note onset detection using one semitone filter-bank for MIREX 2009. In *Proceedings of the Music Information Retrieval Evaluation Exchange - MIREX 2009, Onset Detection Contest* (pp. 84, 95-96). Kobe: International Society for Music Information Retrieval. Retrieved July 14, 2016, from <http://grfia.dlsi.ua.es/repositori/grfia/pubs/238/Pl.pdf>
- Pöhlmann, B. (1981). Stimmbildung. In S. Bimberg, G. Fredrich, F. Höft, H. Irrgang, & C. Neumann (Eds.), *Handbuch der Chorleitung* (pp. 106-143). Leipzig: Deutscher Verlag für Musik.
- Rapoport, E. (1996). Emotional expression code in opera and lied singing. *Journal of New Music Research*, 25 (2), 109-149. doi: 10.1080/09298219608570700
- Rink, J. (2002). Analysis and (or?) performance. In J. Rink (Ed.), *Musical performance: A guide to understanding* (pp. 35-58). New York: Cambridge University Press.
- Scherer, K.R., Sundberg, J., Tamarit, L., & Salomão, G.L. (2015). Comparing the acoustic expression of emotion in the speaking and the singing voice. *Computer Speech and Language*, 29, 218-235. doi: 10.1016/j.csl.2013.10.002
- Sieglwart, H., & Scherer, K.R. (1995). Acoustic concomitants of emotional expression in operatic singing: The case of Lucia in *Ardi gli incensi*. *Journal of Voice*, 9 (3), 249-260. doi: 10.1016/S0892-1997(05)80232-2
- Smith, B., & Sataloff, R.T. (2006). *Choral pedagogy* (2nd ed.). San Diego: Plural Publishing Inc.
- Sundberg, J. (1987). *The science of the singing voice*. Dekalb, Illinois: Northern Illinois University Press.
- Sundberg, J. (2000). Emotive transforms. *Phonetica*, 57, 95-112. doi: 10.1159/000028465
- Sundberg, J., Iwarsson, J., & Hagegård, H. (1994). A singer's expression of emotions in sung performance. *STL-QPSR*, 35 (2-3), 81-92. Retrieved June 20, 2016, from http://www.speech.kth.se/prod/publications/files/qpsr/1994/1994_35_2-3_081-092.pdf
- Sundberg, J., Lã, F.M.B., & Himonides, E. (2013). Intonation and expressivity: A single case study of classical Western singing. *Journal of Voice*, 27 (3), 391.e1-391.e8. doi: 10.1016/j.jvoice.2012.11.009
- Thomas, K. (1979). *Lehrbuch der Chorleitung*. Band I (19th ed.). Wiesbaden: Breitkopf & Härtel.
- Timmers, R. (2007). Vocal expression in recorded performances of Schubert songs. *Musicae Scientiae*, 11 (2), 237-268. doi: 10.1177/102986490701100205
- Trojan, F. (1952). *Der Ausdruck der Sprechstimme: eine phonetische Lautstilistik*. Wien, Düsseldorf: Verlag für medizinische Wissenschaften Wilhelm Maudrich.

DISCOGRAPHY

- Bernius, F. (2006). An die Sonne [Recorded by Kammerchor Stuttgart, A. Rothkopf (piano), F. Bernius (conducting)]. On *Franz Schubert, Chorwerke* [CD]. PH06020. Neuhausen: Profil Medien. (Date of recording: 1987)
- Ehmann, W. (n.d.). Jesu, meine Freude [Recorded by Westfälische Kantorei, H. Flebbe, R. Pax, F. Haasemann, W. Kastrup, J. Kortendiek, A. Schönstedt, W. Ehmann (conducting)]. On *Bach, Motets BWV 226-228* [m4a]. Bibliothèque nationale de France. (Date of recording: 1958)
- Ericson, E. (2012). Jesu, meine Freude [Recorded by Eric Ericson Chamber Choir, Drottningholm Baroque Ensemble, E. Ericson (conducting)]. On *Bach Motets BWV 225-230* [CD]. 5099960230. Estocolmo: EMI Classics. (Date of recording: 1991)
- Gardiner, J.E. (1997). An die Sonne [Recorded by Monteverdi Choir, M. Bilson (forte-piano), J.E. Gardiner (conducting)]. On *Schubert, Choral Works* [CD]. 4544282. Londres: Philips Classics. (Date of recording: 1994)
- Gardiner, J.E. (2012). Jesu, meine Freude [Recorded by Monteverdi Choir, O. Reimers (cello), V. Botwright (double bass), G. Farkas (bassoon), J. Johnstone (organ), J.E. Gardiner (conducting)]. On *Bach Motets* [CD]. SDG716. Londres: Soli Deo Gloria. (Date of recording: 2011)
- Herreweghe, P. (2011). Jesu, meine Freude [Recorded by Collegium Vocale Gent, A. Zweistra (cello), M. Shalinsky (double bass), M. Gratton (organ), P. Herreweghe (conducting)]. On *Johann Sebastian Bach, Motetten* [CD]. LPH002. EU: Phi.
- Neumann, P. (1997). An die Sonne [Recorded by Kölner Kammerchor, E. Schneider (piano), P. Neumann (conducting)]. On *Franz Schubert, An die Sonne* [CD]. 83.138. Stuttgart: Carus-Verlag. (Date of recording: 1996)
- Rilling, H. (1999). Jesu, meine Freude [Recorded by Gächinger Kantorei, Bach-Collegium Stuttgart, H. Rilling (conducting)]. On *J. S. Bach, Motets* [CD]. 92069. Holzgerlingen: Hänssler Classic. (Date of recording: 1990)
- Straube, J. (1997). An die Sonne [Recorded by Norddeutscher Figuralchor, A. Staier (Hammerflügel), J. Straube (conducting)]. On *Franz Schubert, Werke für Chor und Klavier* [CD]. CTH 2358. Baden: Thorofon, Bella Musica Edition.
- Thomas, K. (2004). Jesu, meine Freude [Recorded by Thomanerchor Leipzig, Gewandhausorchester Leipzig, K. Thomas (conducting)]. On *Johann Sebastian Bach, Motetten BWV 225-230* [CD]. 0013292BC. Hamburg: Edel Classics. (Date of recording: 1958/1959)

CAITI HAUCK is Research Associate at the Institute of Musicology of the University of Bern (Switzerland). She holds a PhD in Music, a Masters in Music, and a Bachelor in Music Education from the University of São Paulo (Brazil). Her PhD research discussed aspects of text diction in choral works sung in German and was partly conducted at the University of Cambridge (UK), with the support of a CAPES Foundation grant. Her current research interests include the 19th-century choral movement and musical performance studies. Besides her research activity, Caiti works as a double bassist and conductor of the men's choir Chorale de la Police de Lausanne.