TEXT ARTICULATION AND MUSICAL ARTICULATION IN CHORAL PERFORMANCE: A CASE STUDY

Caiti Hauck

Caiti Hauck is Marie Curie Fellow at the Institute of Musicology of the University of Bern. She conducts the research project "CLEFNI – The choral life in the cities of Bern and Fribourg in the long nineteenth century", which is funded by the European Union's Horizon 2020 Research and Innovation Programme through the Marie Skłodowska-Curie Actions (MSCA).

Caiti holds a PhD in Musicology (2017), a Master of Arts in Music (2012), and a Bachelor in Music Education (2009) from the University of São Paulo (Brazil). Her PhD research was partly conducted at the University of Cambridge (UK) and was funded by the CAPES Foundation (Ministry of Education of Brazil). In addition to her research activity, she conducts the men's choir of the Lausanne Police.

Abstract: Text articulation and musical articulation in choral performance: a case study

In music, the term 'articulation' is often understood as the relationship between contiguous notes in terms of connecting them or not. In vocal music, 'articulation' also means the physical action of producing vowels and consonants. Is there a relationship between these two notions of 'articulation'?

This article discusses the influence of text articulation on musical articulation in choral performances by comparing and contrasting analyses of recordings with bibliographical studies and interviews with conductors. It discusses the effects of different durations and dynamics of vowels and consonants on legato and non-legato articulations. Results suggest that the manner of articulating the text is intrinsically related to the resulting musical articulation.

Text articulation and musical articulation in choral performance: a case study

In music, the term 'articulation' is often understood as the relationship between contiguous notes in terms of connecting them or not. In *Grove Music Online*, for instance, Chew explains that 'the term "articulation" refers primarily to the degree to which a performer detaches individual notes from one another in practice (e.g. in staccato and legato)'.[1] In vocal music, 'articulation' may also refer to the formation of vowels and consonants, as described by Garretson: '*Articulation* pertains to the physical action of the articulating organs (tongue, lips, teeth, palate, and lower jaw) in forming and altering the channels and in projecting the various vocal sounds necessary to achieve intelligible communication'.[2] Such definitions, however adequate in themselves, do not reveal that text articulation and musical articulation are related. For that matter, neither does musical notation.

Western musical notation is relatively precise when indicating the pitch and the duration of notes; however, its system does not provide for elements such as dynamics, timbre, or text diction to be notated with the same degree of precision. With regard to the sung text, musical notation shows the duration of a syllable as a whole, not specifying for how long and how loudly each vowel or consonant should be sung. As a result, notation makes the relationship between text articulation and musical articulation potentially invisible. Yet the use of the same term – 'articulation' – does suggest a connection. Indeed, authors on both vocal pedagogy and choral singing mention, for instance, that the manner of articulating the text has an effect on the legato line, as I will describe later. Nonetheless, the impact of text articulation on musical articulation is an element that has not been extensively investigated in performance.[3] What exactly constitutes the influence of text articulation on musical articulation? What are the effects of different durations and dynamics of vowels and consonants on musical articulation?

In this article, I discuss the influence of text articulation on musical articulation in choral performances of works sung in German. For this purpose, I used a mixed-methods approach. Firstly, I analysed writings on choral conducting, on vocal pedagogy, and on German diction for singing. These bibliographical studies were aimed at identifying what conductors and singers have to say about the effect of text articulation on musical articulation – for example by means of the duration and the dynamics (i.e. loudness) of vowels and consonants. Secondly, I interviewed three conductors: Stephen Cleobury, Martin Ennis, and Peter Neumann. Thirdly, I analysed the pieces 'Abendständchen' and 'Darthulas Grabesgesang' from the *Drei Gesänge für sechsstimmigen Chor a cappella* op. 42 by Johannes Brahms, which were used as case studies. These case studies consisted of analyses of the score and, principally, of recorded performances of these pieces. These were made using the software Sonic Visualiser[4] and also by listening critically. The analyses of the recordings were aimed at identifying possible influences of text articulation upon musical articulation, mainly by measuring the duration and the dynamics of vowels and consonants.

In the first part of this article, I present the results of the bibliographical studies and the interviews. In the second part, I present the results of the analyses of recordings. Finally, I present my conclusions drawn from looking synthetically at the data collected in the bibliographical studies, in the interviews, and in the case studies.

Bibliographical studies and interviews

In order for the text to be intelligible, some conductors advise that the choir exaggerate the pronunciation – that is, the articulation – of the text. Bastian and Fischer recommend 'that

singers not only accurately pronounce the text to be sung, but also articulate it […] intensively'.[5] Halsey states: 'When singers have the feeling that they exaggerate, the text may still be indistinct to the public!'.[6] And, in comparison to *a cappella* singing, Neumann explains that '[w]hen the orchestra joins in […], the clarity of the consonants must be further enhanced, because the overtones of the strings, in particular, absorb much of the articulation noise'.[7] Likewise, Ehmann recommends that the pronunciation be exaggerated, yet he warns against an overly incisive articulation, for this would interrupt the sound flow:

> In choral singing, every singer must *exaggerate the clarity of pronunciation.* For the listener, the exaggeration is then attenuated to normal naturalness by the mass of the choir and the size of the room. However, beware of *too sharp an articulation*, which turns our speech tools into cutting tools; it cuts the words and constantly interrupts the flow of the sound. The plosives (*p*, *t*), in particular, should not be too sharp.[8]

Ehmann is pointing to a conflict between clear articulation and legato line. Cleobury explains that balancing these two elements is precisely one of the technical skills a singer needs to develop:

> There is always the dichotomy between, on the one hand, achieving clarity of enunciation, clarity of the delivery of the text, over against the desire of the singer to produce a cantabile, legato line. And those two things actually stand to some degree in opposition. So the really skilful singer has to find a way to deal with them.[9]

Essentially, both Cleobury and Ehmann are implying that the articulation of the text has an impact on legato: articulating the text – especially consonants – too vigorously could result in a musical articulation other than legato. As a matter of fact, Ehmann and Haasemann argue that, in a strict sense, legato in singing would happen only when there are no consonants:

> Legato means bound together and is represented by a slur. In string instruments, this means that all notes under a slur are to be played in one bow; in the case of wind instruments, it means that all notes under a slur should not be initiated again with the tongue. Consequently, in the narrower sense the word legato and the legato slur should be used in singing when several notes are to be sung on one syllable (*vocalise*). However, in general usage in singing, the word legato with its slur is also used when the notes are to be sustained, and several syllables need to be tightly and softly connected with each other – or even within each other.[10]

The influence of text articulation on musical articulation is mentioned by different conductors and singers. Their suggestions on how to deal with this is described below.

As a general rule for achieving legato, Ehmann and Haasemann advise that vowels should be long and consonants 'short and resonant'.[11] The use of long vowels in legato is also recommended by Hammar.[12] Several authors explain that, in legato, a vowel should have a steady dynamic for the entirety of its duration. Miller states that '[t]he vowel should continue onward at the same degree of intensity until it reaches the consonant, unless the composer has indicated a dynamic shading, or unless the overall phrase shape requires dynamic change'.[13] Similarly, Johnston affirms: 'All vowels must maintain the same degree of intensity until the next phoneme takes over'.[14] Hammar believes that '[o]ne of the most neglected aspects of legato singing is maintaining the core of the vowel sound through the entire length of the note and into the consonant'.[15] A decrescendo while singing a long vowel is also described by Ehmann and Haasemann as one of the elements that disturbs legato.[16]

With regard to consonants, some conductors explain that their characteristics – whether a plosive or a voiced consonant, for example – may variously facilitate a more detached or more

legato articulation. Halsey advises conductors to use each type of consonant as a tool to achieve different musical articulations. Voiced consonants, he states, 'can help carrying legato, and do not need to interrupt the vocal flow. […] Conversely, voiceless consonants can be used as a driving force, they energise the next vowel'.[17] In the interview I carried out with Ennis, he illustrates his considerations concerning consonants and musical articulation when dealing with Baroque music: 'If you make a *k* in the middle of a word, the sound stops momentarily. So, should lots of consonants in a text give rise to a slightly more aerated, slightly more articulated style?'.[18] According to Swan, different consonantal sounds may indicate the type of musical articulation of an excerpt:

> Examine the consonants in the text to determine the composer's concepts of articulation of a specific phrase. Consider how consonants carrying pitch (*m*, *n*, *l*, etc.) influence a more legato phrasing, and how explosive consonants (*p*, *t*, *k*, etc.) contribute to a nonlegato articulation.[19]

Nonetheless, most authors refer to the manner of producing consonants – as opposed to the fact that they be voiced or voiceless – when relating them to musical articulation. Miller, for example, states that '[h]ow voiced and unvoiced consonants are handled has a direct effect on legato, word inflection, phrase shape, dynamic intensity and control, and, most important, interpretation'.[20]

In legato, Ehmann and Haasemann propose that, at the beginning of a word or syllable, the German voiced consonants *l*, *m*, *n*, *ng* ([ŋ]), *w* ([v]), and *s* ([z]) should be longer than usual.[21] The consonants should be anticipated[22] and should occupy a fourth of the duration of the preceding note (here, Ehmann and Haasemann acknowledge one exception to their general rule of using short consonants in legato, quoted earlier). Garretson explains that the voiced consonants *m*, *n*, and *l* 'have a certain degree of sustaining power. As such they may be effectively used, especially in legato singing, to bridge more smoothly the gap between the various vowel sounds'.[23] He proposes that the *m* of the word 'amen' be lengthened to enhance legato; this consonant should be anticipated and occupy a fourth of the duration of the vowel *a* (p. 103). Hammar states that the consonants *m*, *n*, and *ng* ([ŋ]) 'must retain a vital humming sound' and also suggests the lengthening of the *m* of 'amen'.[24]

As regards voiceless consonants, especially the plosives *p*, *t*, and *k*, different authors have opposed points of view. Hammar affirms that, in legato, '[t]he hard consonants [i.e. plosives] must be softened so as not to interfere with the smooth flow of the music'.[25] Garretson recommends that 'the explosive qualities of the consonants should be minimized if they are to be smoothly blended with the vowel sounds'.[26] These authors do not describe how to 'soften' a consonant, yet I suppose this may mean that voiceless plosives (in English and German, for instance)[27] should be pronounced without aspiration.

Johnston has an opposite opinion. She argues that the plosive (and aspirated) consonants of German and English do not prevent but, rather, assist legato:

> It is a common but unfounded fear that the aspirated release of a final consonant results in a disturbance of the legato line. This could not be further from the truth. In reality, an aspirated consonant fills the gap between the words, resulting in greater legato […]. The vocal line would be broken only if there were an inserted space or rest. Further, one must remember that not only the vowels carries the tone. The sound of a consonant still continues the line, be it voiced or voiceless.[28]

A similar opinion is expressed by Marshall, with regard to consonants in English:

Some singers are reluctant to pronounce a well-aspirated voiceless consonant for fear of 'breaking the line.' […] 'The line' is broken whenever there is a gap in it; and the neglect or omission of a voiceless consonant leaves a meaningless vacuum in the flow of the music as well as of the words. Far from 'breaking the line,' clearly articulated consonants help not only to express the full values of the song but also to focus and project the singer's tone.[29]

Ehmann and Haasemann acknowledge that some German consonants – specifically the plosives *b*, *d*, *g*, *p*, *t*, and *k*, and the voiceless fricatives *f* (and *v*, also pronounced as [f]), *sch* ([ʃ]), *ss* and *ß* ([s]), *ch* ([x] as in 'ach', [ç] as in 'ich') – are 'difficult to place' in the legato line. They explain that 'the faster, more accurately, and more flexibly the singer pronounces the consonants […], the denser the legato (cantabile) becomes'.[30] As quoted earlier, these authors propose that, in legato, consonants be short, but not softened.

Ericson, Ohlin, and Spångberg differentiate a dense legato from lighter legato styles precisely in terms of the duration of vowels and consonants. Speaking in relation to lighter styles, they suggest that '[t]he difference in character […] lies primarily in the treatment of the text, specifically, using shorter vowels and more vigorous consonants[31]'.[32] In songs or passages in staccato, they advise: 'Use short vowels and fast consonants for agile singing' (p. 47). In the interview I did with Cleobury, he explains that, 'if you are singing marcato, [for example in a phrase including the word] "mighty" the *m* has got to be done quite quickly'.

Some authors, however, imply that shorter vowels followed by longer (and not faster) consonants may also lead to non-legato articulations. Bastian and Fischer explain that, in portato or staccato passages 'in which the choir imitates instrumental articulations', vowels should be shortened and rapidly give place to the following consonant – for instance, when singing the *l* of the word 'alles' or the *k* of 'Frohlocket' in the chorus 'Komm, holder Lenz' (of Haydn's *Jahreszeiten*), as well as when the choir sings 'don don don' in madrigals or folk songs.[33] This would result in lengthening the consonant, which, in Bastian and Fischer's example, includes not only the voiced *l* and *n*, but also the voiceless plosive *k*.[34]

In order to achieve a dance-like articulation in Bach's works, the conductor Richard Marlow also proposes that consonants be lengthened (which implies that the vowel will be shortened):

Words can be very important in helping create the style. They obviously help color the thing but I work very hard on the soft consonants – the *l*'s and the *n*'s. So often we just hear the vowel going through. But if it's a vowel followed by an *l* or an *n* or something, you can very often get a wonderful resonance – not just making the text audible but a wonderful phrasing. In, say, *Singet dem Herrn* [BWV 225 by Bach], to make it rather bouncy, you don't have to detach it by singing 'Sing-et' but if you sing the *ng*, – 'SiNG-et' – you actually have an articulation that springs. […] Words can get a nice spring and dance-like feel into, for example, Bach.[35]

Another example of shortening vowels and lengthening consonants in non-legato articulations is described by Elliott in relation to the opening phrase of the movement 'Wenn kömmst du, mein Heil', of the cantata *Wachet auf, ruft uns die Stimme* BWV 140 by Bach. When singing the word 'kömmst', Elliott advises shortening the vowel *ö* and lengthening the *m* in order to create a contrast to the preceding downbeat, and thus, a lighter articulation.[36]

The data collected in the bibliographical studies and interviews point to a relationship between legato and non-legato articulations and the duration and dynamics of vowels and consonants. The data presented can be summarised as follows:

- Long vowels: legato[37]
- Steady dynamic over a vowel: legato[38]
- Short consonants: legato (general rule)[39]
- Lengthened voiced consonants (especially *m*, but excluding plosives): legato[40]
- Softened plosive consonants: legato (plosives disturb legato)[41]
- Short plosive consonants: legato (plosives do not disturb legato)[42]
- Short plosives and voiceless fricatives: legato[43]
- Short vowels, vigorous consonants: lighter styles[44]
- Short vowels, fast consonants: staccato[45]
- Fast consonants: marcato[46]
- Short vowels, lengthened consonants (mainly voiced consonants, but occasionally also plosives): portato, lighter articulation[47]

Two conflicting recommendations are observed in these data. The first one has already been mentioned: some authors believe that plosive consonants should be softened because they disturb the legato line, while others affirm that they do not disturb legato and simply need to be short. The second relates to non-legato articulations: some authors advise the use of short consonants, while others recommend lengthened consonants. Analyses of recordings may shed some light on these points.

Analyses of recordings

As mentioned earlier, the pieces used as case studies were 'Abendständchen' and 'Darthulas Grabesgesang' from the *Drei Gesänge für sechsstimmigen Chor a cappella* op. 42 by Brahms. Recordings of 'Abendständchen' were analysed by means of critical listening, which aimed at identifying how voiceless plosive consonants were pronounced in a legato piece. Recordings of 'Darthulas Grabesgesang' were analysed using the software Sonic Visualiser[48]; these analyses aimed at identifying how different durations and dynamics of vowels and consonants relate to musical articulation. Durations of vowels and consonants were measured via spectrogram visualisations and critical listening, aided by the plugins Note Onset Detector and Onset Detection Function.[49] Durations of consonants were measured from the beginning of phonation until the moment the consonant is no longer audible. Durations of vowels were measured from the end of the consonantal sound (rather than the complete vowel formation) to the beginning of the next consonant. Dynamic analyses (i.e. analyses of the loudness of vowels and consonants) were based on the information provided by the plugin Loudness.[50]

The recordings of 'Abendständchen' used in the analyses were those directed by Frieder Bernius,[51] Marcus Creed,[52] Eric Ericson,[53] John Eliot Gardiner,[54] Peter Neumann,[55] and Robert Shaw.[56] Recordings of 'Darthulas Grabesgesang' were directed by Frieder Bernius, Marcus Creed, Eric Ericson, John Eliot Gardiner, and Peter Neumann.[57]

'Abendständchen'

'Abendständchen' is the first song of the *Drei Gesänge* op. 42 by Brahms, set to a poem by Clemens Brentano. Musicality and synaesthesia are hallmarks of Brentano's poem, which makes use of various vowel repetitions (e.g. *ö*, *a*) and cross-references the five senses (e.g. sight and hearing). Its content is both wistful and hopeful. Brahms' composition is strongly related to the text and its tranquil character suggests a legato articulation throughout the piece. Such characteristics could be a reason to soften consonants in the performance of this piece. However, analyses of recordings do not corroborate this idea. The plosives *t* and *k* are virtually

always aspirated – occasionally clearly aspirated –, for instance in the words 'Flöte', 'Töne', 'stille', 'Bitten', or 'kühlen'. Exceptions are the *t* of 'Bitten' in Gardiner's recording, which is not aspirated, and the *k* of 'klagt',[58] which is aspirated only in Gardiner's recording. Special cases occur when a word ends in a plosive and the following word also starts with a plosive, as in 'klagt die', 'Nacht die', and 'blickt zu' (the *z* is pronounced as [ts]). The *t* and *d* of 'klagt die' are clearly articulated (and aspirated, in the case of the *t*) in Shaw's and Gardiner's recordings; in the others, one of the plosives is omitted – mostly the *d*. In 'Nacht die' the *t* is virtually always omitted, with the exception of Gardiner's recording. In 'blickt zu', the final *t* of 'blickt' is omitted in all recordings.

Analyses of recordings thus suggest that conductors generally do not opt for softening consonants in legato. The exceptions are mainly cases when two or more plosives occur at word boundaries.

'Darthulas Grabesgesang'

'Darthulas Grabesgesang' is the third and final song of op. 42 by Brahms, set to a poem by Johann Gottfried Herder, which is based on the alleged translations into English by James Macpherson of the fictional Gaelic poet, Ossian.[59] The poem is a funeral song for Darthula, daughter of Kola, who was killed in a battle. Most parts of Brahms' composition express sorrow for Darthula's death, however the middle section – with tempo indication *Poco animato* – has a lighter and more optimistic character. In this section spring is evoked, and the choir sings: 'Wach auf, wach auf, Darthula!' ('Wake up, wake up, Darthula!').

Figure 1 Bars 48–51 of 'Darthulas Grabesgesang' by Brahms

In bar 48, the first of the *Poco animato* section, there are no indications related to musical articulation, as illustrated in Figure 1. Yet, in the recordings I analysed, different musical

articulations are heard: in some recordings this bar sounds quite legato, while others have a lighter articulation, even a portato. Here, the choir sings the words 'wach auf wach', each one with the duration of a crotchet. I analysed the duration and the dynamics of the vowels and consonants of these words, aiming to identify how they relate to the different musical articulations.

Table 1 presents the durations of the sounds [v], [ɑ], [x], [ɑo], [fv], [ɑ], [x] of 'wach auf wach' in bar 48.[60] The first column presents the conductor of each recording. The second column presents the mean beats per minute (bpm) of the crotchets in this bar.[61] The following columns alternately present the duration in milliseconds of a vowel or a consonant and the percentage of the duration of the mean bpm that is occupied by it.[62] In order to facilitate comparison and contrast, the table is ordered so as to present the recordings sequentially from those that sound more legato to those that have a portato articulation. The shortest duration of each vowel or consonant is indicated in italics; the longest is in bold. Underlined durations and percentages refer to the longest and the shortest sound in each recording.

Conductor	Mean bpm crotchet	[v]	%	[ɑ]	%	[x]	%	[ɑo]	%	[fv]	%	[ɑ]	%	[x]	%
↑ *Legato*															
Ericson	154	**214**	**55.0**	175	45.0	137	35.2	**302**	**77.6**	*127*	*32.6*	173	44.5	151	38.8
Neumann	168	169	47.3	**196**	**54.9**	*116*	*32.5*	232	65.0	139	38.9	**197**	**55.2**	*104*	*29.1*
Creed	156	173	45.1	131	34.1	140	36.5	262	68.2	185	48.2	*127*	*33.1*	162	42.2
Bernius	160	098	26.1	150	40.0	186	49.6	214	57.1	162	43.2	144	38.4	185	49.3
Gardiner	192	*055*	*17.6*	094	30.1	**174**	**55.8**	*149*	*47.8*	152	48.7	116	37.2	**174**	**55.8**
↓ *Portato*															

Table 1 Durations of the sounds of 'wach auf wach' in bar 48 of 'Darthulas Grabesgesang' by Brahms

The data regarding the dynamics of vowels and consonants are illustrated in the spectrograms of this passage. Figures 2 to 6 reproduce these spectrograms and refer to the recordings directed by Ericson, Neumann, Creed, Bernius, and Gardiner, respectively. Dynamics are indicated by the yellow line created by the plugin Loudness. The sung text and its duration in milliseconds are indicated in the labels in red. White vertical lines demarcate vowels and consonants.
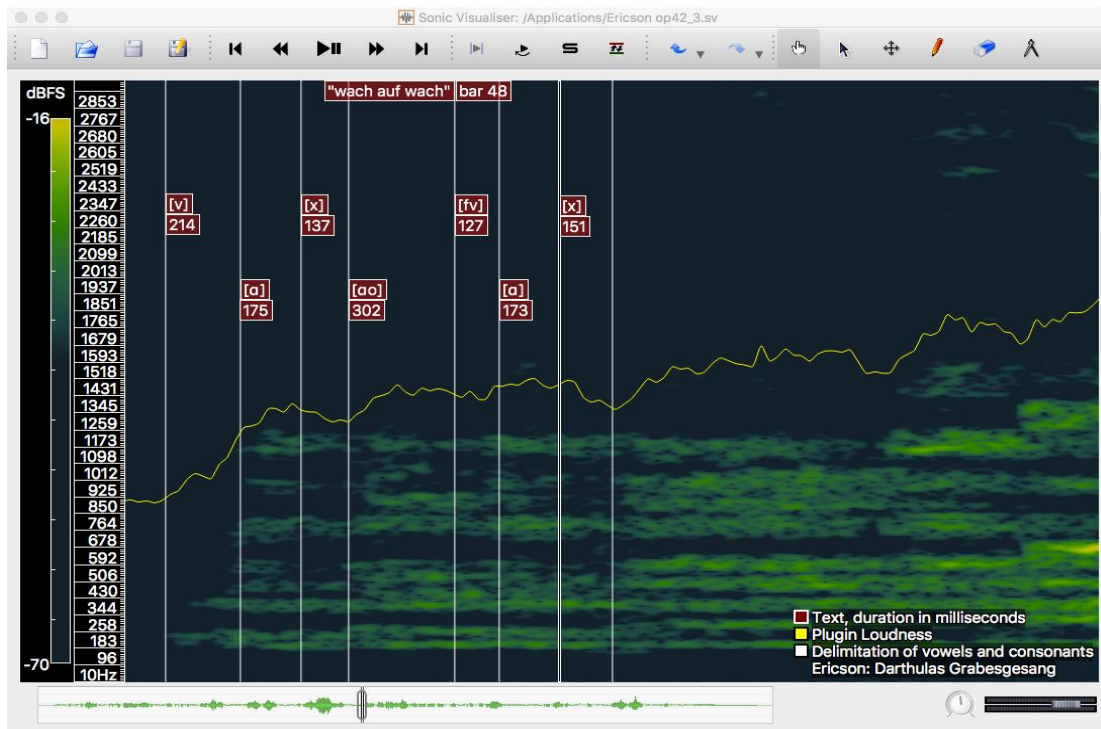
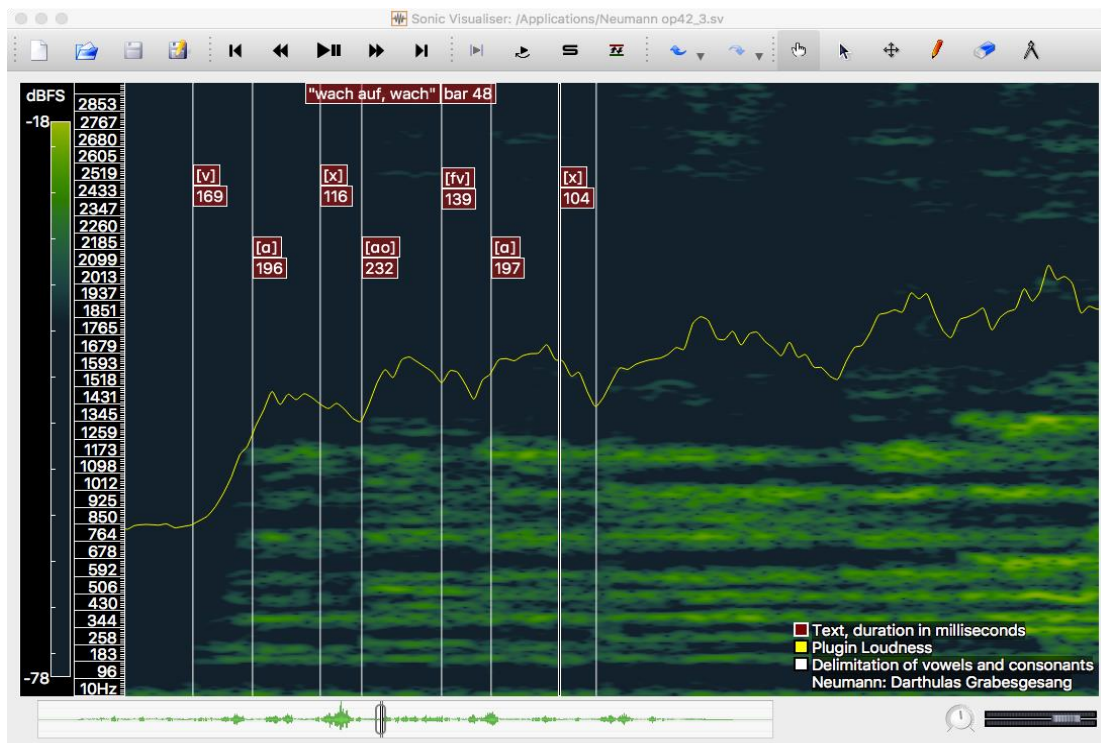**Figure 2** Spectrogram of the recording directed by Ericson of 'Darthulas Grabesgesang' by Brahms

Figure 3 Spectrogram of the recording directed by Neumann of 'Darthulas Grabesgesang' by Brahms
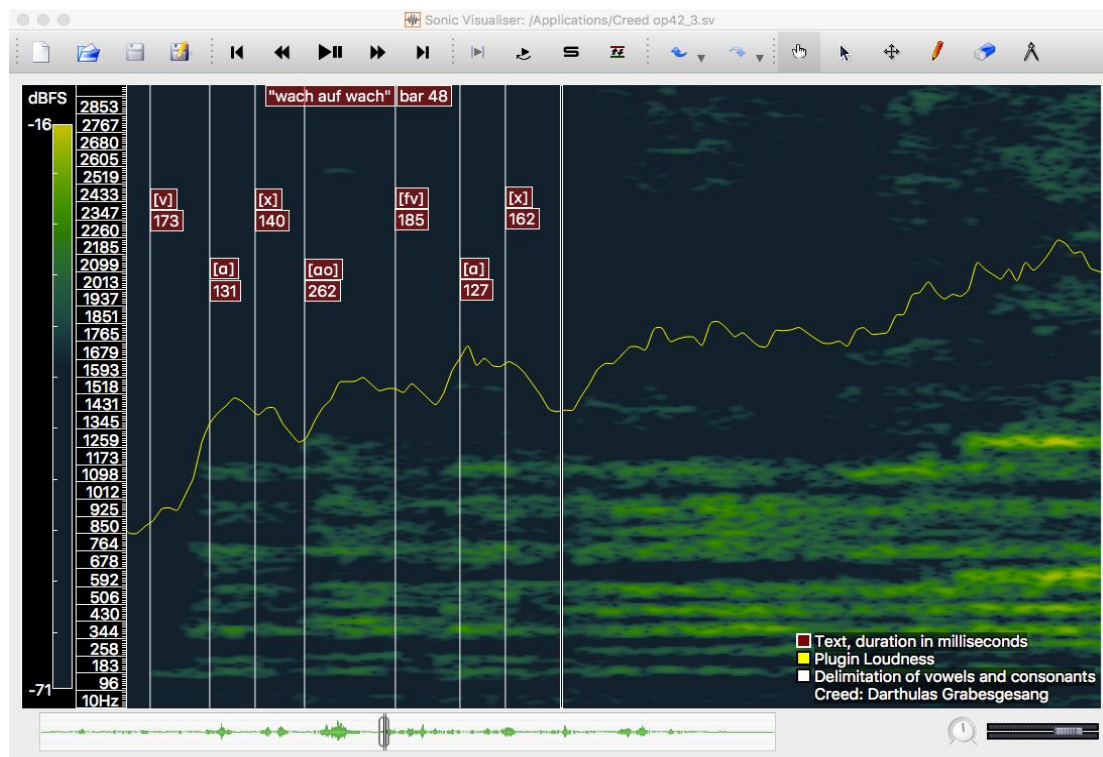
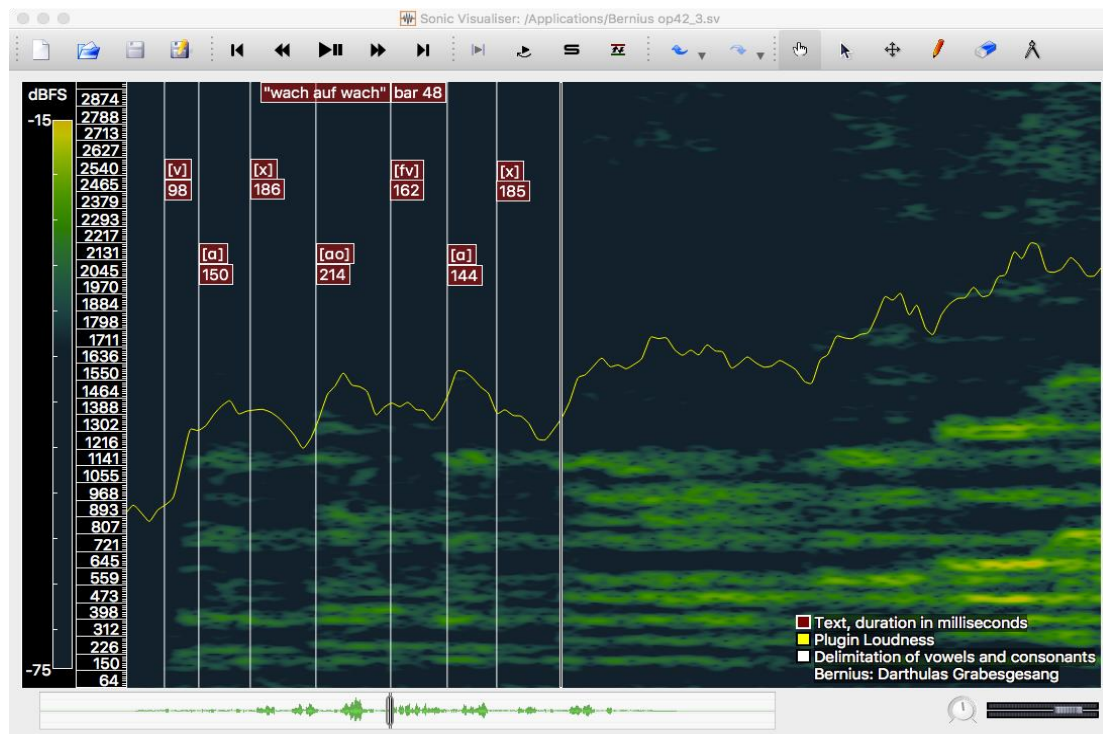**Figure 4** Spectrogram of the recording directed by Creed of 'Darthulas Grabesgesang' by Brahms

Figure 5 Spectrogram of the recording directed by Bernius of 'Darthulas Grabesgesang' by Brahms
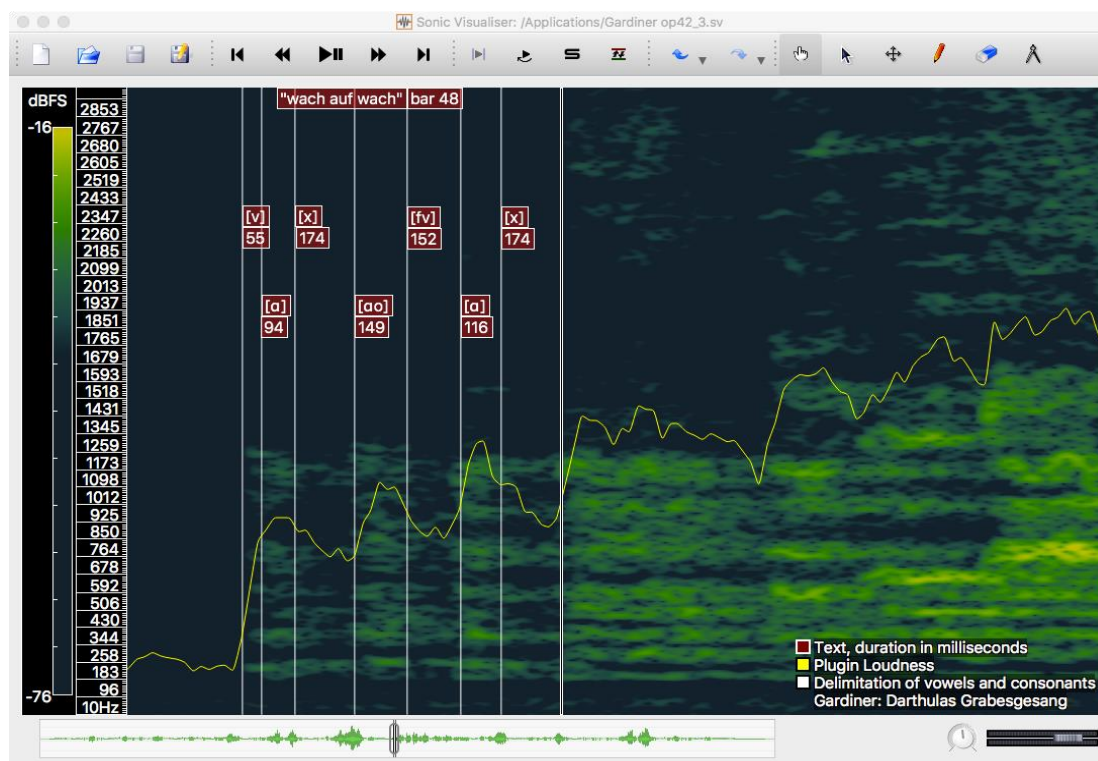
Figure 6 Spectrogram of the recording directed by Gardiner of 'Darthulas Grabesgesang' by Brahms

Of these recordings, the one that sounds the most legato is Ericson's. The dynamic contour created by the plugin Loudness indicates that the overall dynamic of bar 48 is quite uniform. The initial consonant [v] is the longest of all five recordings, occupying 55% of the mean bpm of this passage, and its dynamic increases slowly. The diphthong [ɑo] is also the longest of the analysed recordings and occupies 78% of the mean bpm; it has a slight crescendo at the beginning, but the dynamic remains steady at the transition to the next consonant. The consonant cluster [fv] is the shortest of all the recordings, with 33% of the mean bpm. Apart from the initial [v], all consonants are shorter than vowels, and their dynamics are just slightly softer than the dynamics of vowels.

Neumann's recording also sounds legato, but somewhat less than Ericson's. As in the preceding recording, the initial [v] is long with a gradual dynamic increase. Both vowels [ɑ] are the longest of all the recordings, occupying 55% of the mean bpm, and both consonants [x] are the shortest, with roughly 30% of the mean bpm. All consonants are shorter than vowels, and dynamics decrease slightly by the end of vowels and consonants (especially the voiceless consonants).

Creed's recording sounds less legato than the ones by Ericson and Neumann. The [v] is shorter than in the preceding recordings, yet it still occupies 45% of the mean bpm. The second 'wach' is accentuated; this is reflected by the long duration of [fv] and by the fast increase in dynamic from the end of this consonant cluster (that is, from the articulation of the [v]) into the vowel [ɑ]. Yet this second vowel [ɑ] is the shortest of all the recordings, with 33% of the mean bpm. Consonants are mostly longer than vowels; the only exception is the diphthong [ɑo]. The dynamic contour indicates greater variations than in Ericson's and Neumann's recordings.

Bernius' recording has a lighter articulation; here, proportions begin to change. The initial [v] is very short, occupying 26% of the mean bpm, and its dynamic increases rapidly. All other

consonants are longer than vowels, with the exception of the diphthong [ɑo]. The dynamic variation between vowels and consonants is greater than in Creed's recording, and the dynamic contour indicates that the dynamics of vowels decrease before they achieve the next consonant.

Gardiner's recording is the one which has the more detached articulation, sounding like a portato. The initial [v] is shorter than in Bernius' recording, occupying 18% of the mean bpm, and its dynamic increases immediately. This is the only short consonant in the passage: all other consonants are longer than the vowels. They are also the longest out of all the recordings: the cluster [fv] occupies 49% of the mean bpm, and the consonants [x], with 56% of the mean bpm, are the longest sounds within this recording as well. These consonants are even longer than the diphthong [ɑo], which is the longest sound in all the other recordings. The dynamic contour indicates that dynamics decrease significantly when consonants are articulated and before the end of the vowel as well.

The most contrasted recordings are the ones by Ericson and Gardiner. In Ericson's recording, vowels are longer than consonants, with the exception of the initial [v], which is longer than both vowels [ɑ] (but not as long as the diphthong [ɑo]). In Gardiner's recording, the opposite is true: with the exception of the initial [v], all vowels (including the diphthong [ɑo]) are shorter than consonants. These recordings also have the most dissimilar tempi: Ericson's is the slowest and Gardiner's the fastest. Nonetheless, the relation between the durations of vowels and consonants and a more or less legato articulation is also observed in recordings that have similar tempi to one another. Neumann's recording, for instance, sounds more legato than Bernius', even though it is just slightly faster. In Neumann's recording, all vowels are longer than consonants, while in Bernius' – apart from [v] and [ɑo] – vowels are shorter than consonants.

Analyses of recordings indicate that the dynamics of vowels and consonants also play a role in musical articulation. The dynamic contours of Ericson's and Neumann's recordings are much steadier than those of Bernius and Gardiner. The accent on the second 'wach' in Creed's recording contributes to a less legato articulation as well.

In this short passage from Brahms' composition there is little variety of consonantal sounds – just one voiced fricative and two voiceless fricatives. Nor do the recordings exemplify a great diversity of musical articulations, their range being roughly from a legato to a portato. Still, these analyses do bring new elements to understanding of how the duration and the dynamics of vowels and consonants affect musical articulation.

Discussion

The data collected in the bibliographic studies, in the interviews, and in the analyses of recordings show a number of similarities. However, results of the recording analyses reveal various elements about the relationship between text articulation and musical articulation that are not mentioned by conductors in the interviews and in the literature. Moreover, analyses of recordings may shed some light on the conflicting recommendations observed in the bibliographic studies and in the interviews.[63]

Long vowels in legato singing are proposed by Ehmann and Haasemann and by Hammar. Analyses of the recordings of 'Darthulas Grabesgesang' directed by Ericson and by Neumann corroborate this suggestion. In bar 48, Ericson's and Neumann's recordings are the ones that sound more legato, and their long vowels occupy on average 57% of the mean bpm.

The use of short vowels in lighter articulations, as proposed by Bastian and Fischer, Elliott, and Ericson, Ohlin, and Spånberg (and implied by Marlow), is also corroborated by the analyses of recordings. Gardiner's recording of bar 48 of 'Darthulas Grabesgesang' sounds like a portato and, on average, its short vowels occupy 38% of the mean bpm. Vowels in Bernius'

and Creed's recording – which have a lighter articulation than Ericson's and Neumann's ones – occupy on average 45% of the mean bpm.

A steady dynamic over a vowel is recommended by Ehmann and Haasemann, Hammar, Johnston, and Miller, and this is supported by the analyses of 'Darthulas Grabesgesang'. In Ericson's recording, the dynamic contour created by the plugin Loudness reveals that vowels maintain their dynamics until the next consonant is reached. By contrast, Neumann's recording shows, for instance, both a crescendo at the beginning and a decrescendo by the end of the diphthong [ɑo]. Dynamics of consonants in legato are not mentioned in the literature I analysed, yet the dynamic contour of Ericson's recordings is quite uniform and indicates that consonants are barely softer than vowels. The dynamic contour of Neumann's recording is not as uniform, and consonants are always softer than vowels. This may explain why Neumann's recording sounds less legato than Ericson's, even though Neumann's recording presents, on average, slightly longer vowels than Ericson's (58% of the mean bpm in Neumann's, against 56% in Ericson's).

Dynamics of vowels in non-legato articulations are also not mentioned in the analysed literature. Yet, the recordings of 'Darthulas Grabesgesang' directed by Creed, Bernius, and Gardiner indicate that vowels usually have a peak: the dynamics increase and decrease over their durations. This is best observed in the spectrogram of Gardiner's recording, but those of Creed and Bernius also present the same pattern.

Comparing and contrasting the data regarding consonants in both legato and non-legato articulations reveals many specificities. A first issue to be discussed concerns plosive consonants in legato. In the literature I analysed, Garretson and Hammar recommend that plosives should be softened, while Johnston and Marshall argue that they do not disturb legato. As mentioned earlier, analyses of the recordings of 'Abendständchen' do not corroborate Garretson's and Hammar's recommendation: plosives are nearly always aspirated. This suggests that, in general, conductors do not opt for softening consonants in order to achieve legato.

The use of short voiceless fricatives in legato, especially the German *ch* ([x]) – as proposed by Ehmann and Haasemann – is supported by the analyses of 'Darthulas Grabesgesang'. In Ericson's and Neumann's recordings, the two consonants [x] are short and occupy on average 34% of the mean bpm. The cluster [fv] comprises both a voiceless and a voiced fricative, yet it is also short in these recordings, occupying on average 36% of the mean bpm. In legato, Ehmann and Haasemann also recommend that plosives (both voiced and voiceless) be short. The passage of 'Darthulas Grabesgesang' that was analysed does not have plosives, and it was not the aim of this study to measure the duration of plosives in the recordings of 'Abendständchen'; however, it is reasonable to believe that, if plosives are long, they will not enhance legato.

Lengthening voiced consonants (except plosives) is mentioned by Ehmann and Haasemann, Garretson, and Hammar as a means of improving legato. As explained earlier, the passage of 'Darthulas Grabesgesang' used for analysis does not have, for instance, an *m* (which is a consonant mentioned in these three publications). Nonetheless, analyses of recordings do add details to this recommendation and also reveal elements that are not mentioned in the literature. Data regarding the duration of vowels and consonants in the recordings point to a proportional relationship: vowels longer than consonants can enhance legato, while consonants longer than vowels should result in non-legato articulations. Indeed, this proportion is implied in the examples of lengthened voiced consonants in legato provided by Ehmann and Haasemann (German voiced consonants) and Garretson (the *m* of 'amen'): the consonant is lengthened, but it is still shorter than the vowel. Yet, the statement above requires qualification: analyses of recordings also suggest that this proportion can be completely inverted if the

consonant is preceded by a rest (or a short silence), for example at the beginning of a musical phrase. In legato, consonants not preceded by a rest (in the middle of a musical phrase, for instance) may be shorter than the succeeding vowel; but when the consonant is preceded by a rest, it can be longer than the vowel. This is illustrated by Ericson's recording: the first consonant *w* ([v]) is longer than the succeeding vowel and occupies 55% of the mean bpm, but all other consonants are shorter than vowels. The analyses of the dynamics of consonants described above also suggest that, if voiced consonants (except plosives) are lengthened – and especially if they are in the middle of a musical phrase – their dynamics would need to be not much softer than those of vowels, if a legato is to be achieved.

The idea of proportional duration being connected with the position of the consonant also seems valid in the case of non-legato articulations, as suggested by Gardiner's recording of 'Darthulas Grabesgesang'. In this recording, consonants in the middle of the phrase are longer than vowels, but the first consonant *w* ([v]) is shorter than all vowels (and shorter than other consonants as well). Thus, also in non-legato, the proportional duration – consonants longer than vowels – can be inverted when the consonant is preceded by a rest or a silence (at the beginning of a musical phrase, for instance). This may explain the conflict between the recommendation by Ericson, Ohlin, and Spångberg to use fast consonants in staccato and the proposition by Bastian and Fischer, Elliott, and Marlow to use long consonants in non-legato articulations. Ericson, Ohlin, and Spångberg do not provide a specific example to illustrate their statement; however, they refer to staccato, which is the most detached musical articulation.[64] In staccato passages, generally there would be a (short) silence between notes (or syllables). This silence implies that initial consonants (and final ones as well) would most likely be short (either shorter than, or as short as, vowels), otherwise the note (or syllable) would be lengthened, and the silence between notes would not occur (especially if the tempo is fast).

By contrast, the examples of lengthened consonants offered by Bastian and Fischer, Elliott, and Marlow refer to portato-like articulations, and the consonants they cite are not preceded by rests (*l* of 'alles',[65] *k* of 'Frohlocket', *ng* of 'singet', *m* of 'kömmst'). Analyses of recordings cannot corroborate the propositions of these authors, since there is no *l*, *k*, *ng*, or *m* in bar 48 of 'Darthulas Grabesgesang', yet Gardiner's recording does indicate that employing consonants longer than vowels in the middle of a musical phrase can result in a portato articulation. Further, this recording illustrates that voiceless fricatives – and not only *l*, *k*, *ng*, or *m* – can also be lengthened in order to achieve portato. Ultimately, this suggests that lengthening virtually all types of consonants – voiced or unvoiced, plosive, nasal, fricative, or liquid – could be a means of achieving a portato-like articulation.

Conclusion

Results of this research may be summarised as follows:

- In legato, vowels are generally longer than consonants, but a consonant (especially a voiced one) preceded by a rest or a silence can be longer than the succeeding vowel. Plosive consonants are exceptions and should be short (but not necessarily softened). Dynamics between vowels and consonants are mostly constant in legato.

- In portato-like articulation, consonants are generally longer than vowels, but a consonant preceded by a rest or a silence can be shorter than vowels; this could refer to nearly all types of consonants. Dynamics of vowels typically form a peak (increasing and decreasing over the vowel's duration), while dynamics of consonants are generally softer than those of vowels.

These findings indicate that the manner of articulating the text is intrinsically related to the resulting musical articulation. In fact, the results suggest that in vocal performance – or at least in choral performance – one cannot discuss musical articulation without discussing text articulation. Although it is an under-researched element of choral performance, the manner of articulating the sung text has a major effect on performance, notably on musical articulation and, therefore, on performance expressivity.

1   Geoffrey Chew, 'Articulation and Phrasing', *Grove Music Online* (Oxford University Press, 2001) <https://www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000040952> [accessed 12 July 2019].

2   Robert L. Garretson, *Conducting Choral Music*, 8th edn (Upper Saddle River, N.J.: Prentice Hall, 1998), p. 91.

3   Despite the increasing number of studies on musical performance and on performance expressivity, the sung text and its effect on performance remains under researched. For accounts on the role of the sung text in performance expressivity, see Caiti Hauck, 'Diction, Expressivity and Conductor's Choices in Choral Works Sung in German', *Music Performance Research*, 9 (2019), 80–100 <http://www.mpronline.net/Issues/Volume%209%20[2019]/MPR%200129%20Hauck%20(80-100).pdf> [accessed 17 April 2019]; and Daniel Leech-Wilkinson, *The Changing Sound of Music: Approaches to Studying Recorded Musical Performances* (London: CHARM, 2009). <http://www.charm.kcl.ac.uk/studies/chapters/intro.html> [accessed 7 February 2019], especially Chapter 8. As regards musical articulation, Keller argues that 'we can derive principles of musical articulation from language' and draws a parallel between musical articulation and speech articulation. Nonetheless, when writing about singing, he does not consider that the manner of articulating the text may have an effect on the musical articulation. See Hermann Keller, *Phrasing and Articulation: A Contribution to a Rhetoric of Music*, trans. by Leigh Gerdine (London: Barrie & Rockliff, 1966), pp. 31–32.

4   Chris Cannam, Christian Landone, and Mark Sandler, 'Sonic Visualiser: An Open Source Application for Viewing, Analysing, and Annotating Music Audio Files', *Proceedings of the ACM Multimedia 2010 International Conference* (2010), 1467–1468 <http://www.sonicvisualiser.org/sv2010.pdf> [accessed 13 July 2019].

5   Hans Günther Bastian and Wilfried Fischer, *Handbuch der Chorleitung* (Mainz: Schott, 2006), p. 261. '[…] dass die Sängerinnen und Sänger den zu singenden Text nicht nur richtig aussprechen, sondern auch […] intensiv artikulieren'. All translations are by the author.

6   Simon Halsey, *Schott Master Class Chorleitung: Vom Konzept zum Konzert*, written with Wiebke Roloff (Mainz: Schott, 2011), p. 209. 'Wenn die Sänger das Gefühl haben zu übertreiben, ist der text für das Publikum vielleicht immer noch undeutlich!'. All translations are by the author.

7   Peter Neumann, interview with the author, 2015. 'Wenn das Orchester dazu kommt […] muss man die Deutlichkeit der Konsonanten noch verstärken, weil gerade die Obertöne der Streicher viel von den Artikulationsgeräuschen schlucken'. All translations are by the author.

8   Wilhelm Ehmann, *Die Chorführung*, vol 2, 6th edn (Kassel: Bärenreiter, 1981), p. 56, italics in original. 'Beim Chorsingen muß jeder Einzelsänger die *Deutlichkeit der Aussprache übertreiben*. Die Übertriebenheit wird dann durch die Masse des Chores und die Größe des Raumes für die Hörer zur normalen Natürlichkeit abgeschwächt. Man hüte sich jedoch vor einer *zu scharfen Artikulation*, die unsere Sprachwerkzeuge zu Schneidewerkzeugen macht; sie zerschnippelt die Worte und unterbricht ständig den Tonstrom. Besonders die Verschlußlaute dürfen nicht zu scharf genommen werden (p, t)'. All translations are by the author.

9   Stephen Cleobury, interview with the author, 2015.

[10]   Wilhelm Ehmann and Frauke Haasemann, *Handbuch der Chorischen Stimmbildung*, 3rd edn (Kassel: Bärenreiter, 1990), p. 64. 'Legato heißt Bindung und wird durch einen Bogen dargestellt. Dies meint beim Streichinstrument, daß alle unter einem Bogen stehenden Noten auf einen Bogenstrich zu spielen sind; es meint beim Blasinstrument, daß alle unter einem Bogen stehenden Noten nicht erneut mit der Zunge angestoßen werden sollen. Folgerichtig müßten das Wort legato und der legato-Bogen im engeren Sinne beim Singen dann angewendet werden, wenn mehrere Töne auf eine Silbe zu singen sind (Vokalise). Im allgemeinen sängerischen Sprachgebrauch benutzt man aber das Wort legato mit seinem Bogen auch dann, wenn die Töne ganz ausgehalten und mehrere Silben eng und weich aneinander – oder gar ineinander – gehängt werden müssen'. All translations are by the author.

[11]   Ehmann and Haasemann, *Handbuch der Chorischen Stimmbildung*, p. 183. 'kurz und klingend'.

[12]   Russell A. Hammar, *Pragmatic Choral Procedures* (Metuchen, N.J. & London: Scarecrow Press, 1984), p. 190.

[13]   Richard Miller, *On the Art of Singing* (New York [etc.]: Oxford University Press, 1996), pp. 22–23.

[14]   Amanda Johnston, *English and German Diction for Singers: A Comparative Approach* (Lanham [etc.]: Scarecrow Press, 2011), p. 225.

[15]   Hammar, *Pragmatic Choral Procedures*, p. 214.

[16]   Ehmann and Haasemann, *Handbuch der Chorischen Stimmbildung*, p. 64.

[17]   Halsey, *Schott Master Class Chorleitung*, p. 209. 'Sie können helfen, das Legato zu tragen, und müssen den Vokalfluss nicht unterbrechen. […] Die stimmlosen Konsonanten dagegen können als Impulsgeber genutzt werden, sie energetisieren den nächsten Vokal'.

[18]   Martin Ennis, interview with the author, 2015.

[19]   Howard Swan, 'The Development of a Choral Instrument', in *Choral Conducting Symposium*, ed. by Harold A. Decker and Julius Herford (Englewood Cliffs, N.J.: Prentice Hall, 1988), pp. 7–68 (p. 62).

[20]   Richard Miller, *Solutions for Singers: Tools for Performers and Teachers* (Oxford: Oxford University Press, 2004), pp. 116–117.

[21]   Ehmann and Haasemann, *Handbuch der Chorischen Stimmbildung*, pp. 69.

[22]   I have written elsewhere about the anticipation of initial consonants. See Caiti Hauck, 'When are initial consonants articulated in choral performance? Case studies of choral works sung in German', *Music Performance Research* (forthcoming).

[23]   Garretson, *Conducting Choral Music*, p. 103.

[24]   Hammar, *Pragmatic Choral Procedures*, p. 99.

[25]   Hammar, *Pragmatic Choral Procedures*, p. 214.

[26]   Garretson, *Conducting Choral Music*, p. 99.

[27]   English voiceless plosives are 'pronounced with aspiration at the beginning of stressed syllables', see Mehmet Yavaş, *Applied English Phonology*, 2nd edn (Malden, MA: Wiley-Blackwell, 2011), p. 58. German voiceless plosives (*p*, *t*, and *k*), as well as *b*, *d*, and *g* at the end of a word or a syllable, are generally aspirated, see Theodor Siebs, *Deutsche Aussprache: reine und gemässigte Hochlautung mit Aussprachewörterbuch*, ed. by Helmut de Boor, Hugo Moser, and Christian Winkler, 19th rev. edn (Berlin: Walter de Gruyter, 1969), pp. 22 and 104–105.

[28]   Johnston, *English and German Diction for Singers*, p. 227.

[29]   Madeleine Marshall, *The Singer's Manual of English Diction* (New York: Schirmer, 1953), p. 31.

[30]   Ehmann and Haasemann, *Handbuch der Chorischen Stimmbildung*, p. 70. 'schwer einzuordnen'; 'je schneller, genauer, beweglicher der Sänger die Konsonanten […] abspricht, je dichter wird das legato (cantabile)'.

[31]   Ericson, Ohlin, and Spångberg do not describe what is meant by 'vigorous consonants', yet I understand that the consonant should be short with stronger dynamics.

32   Eric Ericson, Gösta Ohlin, and Lennart Spånberg, *Choral Conducting* (Stockholm: Sveriges Körförbunds Förlag, 1976), p. 35.

33   Bastian and Fischer, *Handbuch der Chorleitung*, p. 259. 'in denen der Chor instrumentaler Artikulationsarten nachahmen'.

34   Lengthening plosive consonants means actually lengthening their occlusion, which creates a silence.

35   Jeffrey Sandborg, *English Ways: Conversations with English Choral Conductors* (Chapel Hill, N.C.: Hinshaw Music, Inc., 2001), p. 128.

36   Martha Elliott, *Singing in Style: A Guide to Vocal Performance Practices* (New Haven & London: Yale University Press, 2006), p. 59.

37   Ehmann and Haasemann, *Handbuch der Chorischen Stimmbildung*, p. 183; Hammar, *Pragmatic Choral Procedures*, p. 190.

38   Ehmann and Haasemann, *Handbuch der Chorischen Stimmbildung*, p. 64; Hammar, *Pragmatic Choral Procedures*, p. 214; Johnston, *English and German Diction for Singers*, p. 225; Miller, *On the Art of Singing*, pp. 22–23.

39   Ehmann and Haasemann, *Handbuch der Chorischen Stimmbildung*, p. 183.

40   Ehmann and Haasemann, *Handbuch der Chorischen Stimmbildung*, p. 69; Garretson, *Conducting Choral Music*, p. 103; Hammar, *Pragmatic Choral Procedures*, p. 99.

41   Garretson, *Conducting Choral Music*, p. 99; Hammar, *Pragmatic Choral Procedures*, p. 214.

42   Johnston, *English and German Diction for Singers*, p. 227; Marshall, *The Singer's Manual*, p. 31.

43   Ehmann and Haasemann, *Handbuch der Chorischen Stimmbildung*, p. 70.

44   Ericson, Ohlin, and Spånberg, *Choral Conducting*, p. 35.

45   Ericson, Ohlin, and Spånberg, *Choral Conducting*, p. 47.

46   Cleobury, interview.

47   Bastian and Fischer, *Handbuch der Chorleitung*, p. 259; Elliott, *Singing in Style*, p. 59; Marlow in: Sandborg, *English Ways*, p. 128.

48   See note 4.

49   Chris Duxbury, Juan Pablo Bello, Mike Davies and Mark Sandler, 'Complex domain onset detection for musical signals', *Proceedings of the 6th Digital Audio Effects Workshop (DAFx-03)* (2003) <http://www.eecs.qmul.ac.uk/legacy/dafx03/proceedings/pdfs/dafx81.pdf> [accessed 19 July 2019]. Antonio Pertusa and José M. Iñesta, 'Note onset detection using one semitone filter-bank for MIREX 2009', *Proceedings of the MIREX 2009 - Music Information Retrieval Evaluation eXchange, MIREX Audio Onset Detection,* (2009) <https://grfia.dlsi.ua.es/repositori/grfia/pubs/238/PI.pdf> [accessed 19 July 2019].

50   Jamie Bullock, 'Libxtract: A lightweight library for audio feature extraction', *Proceedings of the International Computer Music Conference* (2007), 25–28 <http://hdl.handle.net/2027/spo.bbp2372.2007.116> [accessed 19 July 2019].

51   Frieder Bernius, *Johannes Brahms, Lieder & Romanzen*, Kammerchor Stuttgart, recorded in Mozartsaal, Liederhalle Stuttgart, on February 1995 (Sony Music, 88875073032, 2015).

52   Marcus Creed, *Johannes Brahms, Secular Choral Songs*, CD1, RIAS Kammerchor, recorded in Jesus-Christus Church, Berlin-Dahlem, on November 1995 and April 1996 (Harmonia Mundi, HMG 501592.93, 2010).

53   Eric Ericson, *Brahms, Zigeunerlieder, Secular Choruses*, Swedish Radio Choir, recorded in Västerled Church, Stockholm, on April 1983 (Teldec, 0630-17426-2, 1997).

54   John Eliot Gardiner, *Brahms, Choral Works*, Monteverdi Choir, recorded in St Giles Cripplegate, London, on November 1990 (Decca, 4757558, 2006).

55   Peter Neumann, *Johannes Brahms, Chorlieder*, Kölner Kammerchor, recorded in Südwest-Tonstudio, Stuttgart, on October 1983 (Carus-Verlag, 83.107, 1987).

[56] Robert Shaw, *Brahms, Liebeslieder Waltzes, Evening Songs*, Robert Shaw Festival Singers, recorded in Church of St. Pierre, Gramat, France, on at August 1992 (Telarc, 80326, 1993).

[57] References to these recordings are described on notes 51 to 55.

[58] German voiceless plosives should be aspirated also when followed by *r* or *l*, see Siebs, *Deutsche Aussprache*, p. 104.

[59] See James Porter, 'Ossian [Oisean, Oisín]', *Grove Music Online*, <https://doi.org/10.1093/gmo/9781561592630.article.47070> [accessed 22 July 2019]; Wolf Gerhard Schmidt, 'Ossian', *MGG Online*, ed. by Laurenz Lütteken, <https://www.mgg-online.com/mgg/stable/16010> [accessed 22 July 2019].

[60] The vowels of the diphthong [ɑo] and the consonants of the cluster [fv] (formed at the boundaries of 'auf' and the second 'wach') were not measured individually, since the aim was to contrast vowels with consonants.

[61] Bpm were calculated with Sonic Visualiser. The mean bpm refers to beats from the third crotchet of bar 48 to the second crotchet of bar 50.

[62] Percentages of the mean bpm are approximate. This happens not only because they were calculated in relation to the mean bpm – rather than the exact bpm of each crotchet of bar 48 –, but also because initial consonants probably occupy a fraction of the preceding beat. (This relates to the anticipation of initial consonants, see note 22.)

[63] Analysis with Sonic Visualiser shows the minute details of a recorded performance and gives them a kind of concreteness that the ear alone might not detect. It is therefore important to bear in mind, when drawing conclusions from them, that such details may not necessarily be the result of conscious and predetermined performance choices by the artists in question, but merely symptoms of the ephemeral and 'accidental' micro-events that occur in any performance. Furthermore, the fact that the analysed recordings are all by respected performers does not remove the possibility that, in some recordings, the relationship between textual and musical articulation might be less effective than in others simply because of the vagaries of the occasion, the location, or the recording technology employed.

[64] Staccatissimo is certainly more detached than staccato, but it is still a type of staccato.

[65] If one chooses to sing staccato the phrase 'bald lebet alles wieder auf' (in the choir 'Komm, holder Lenz', of Haydn's *Jahreszeiten*), I believe the *l* of 'alles' would not be lengthened, since there would be silences between syllables.