journal of interdisciplinary music studies
spring/fall 2013, volume 7, issue 1&2, art. #13071201, pp. 01-28

The rise of fixed pitch systems and the slide of continuous pitch: A note for emotion in music research about portamento

Emery Schubert[1] and Joe Wolfe[2]

[1] School of the Arts and Media, The University of New South Wales
[2] School of Physics, The University of New South Wales

Background in music psychology and history. Psychological and historical factors such as the desire to imitate, and the development of music notation systems and sound recordings, each appear to have contributed to the demise and suppression of *portamento* (pitch bending) in music, and the singing voice in particular. *Portamento* offers potential expressive resources in both music and speech, and so it seems important to examine what has happened to *portamento* in the study of emotion in music – we found limited empirical literature that examines the emotional effects of time-varying *portamento* in music.
Background in music acoustics. The human voice is well suited to producing phonemes and speech, where control of timbre (including vowel sounds) is more important than control of pitch. This physiological and acoustic observation provides us with a clue as to what 'music' may have sounded like if and when it was part of a protolanguage. The physiology and acoustics of the human voice lend it to the production of *portamento* and timbre control better than it does stable pitch, even though fixed pitch tone production is found in much conventional Western music.
Aims. We aim to investigate the nature of *portamento*, and why it has received such little attention in research on emotion in music.
Main contribution. The paper presents evidence that the human voice is well suited to producing *portamento*, and that the near absence of portamento in emotion research study may be because of the difficulty in measuring and analyzing this transient aspect of expression, and an artifact of discrete notation systems, sound recordings and the desire to imitate musical instruments that easily produce fixed pitches, such as the flute and guitar.
Implications. Portamento has known expressive utility in speech, and despite impediments it is important to investigate whether *portamento* has such an expressive capacity in music. Literature on emotion prosody and the psychophysics of *portamento* is drawn on generate hypotheses for future research.

Keywords: *portamento,* music expression, mimcry, music notation, sound recording, emotion, speech, music acoustics, glissando threshold.

•*Correspondence:* Emery Schubert, UNSW, Sydney, 2052, NSW, Australia;
tel: +61-2-9385 6808, e-mail: e.schubert@unsw.edu.au
• *Received:* 31 September 2013; *Revised:* 26 February 2014; *Accepted:* 20 March 2014
• *Available online:* 09 July 2014
• doi: 10.4407/jims.2014.06.001

Introduction

In the compendium of music and emotion edited by Juslin and Sloboda (2010) there is hardly any mention of the influence of *portamento* on emotion in music, despite detailed investigations and summaries of the emotional effects of dozens of other musical features (Juslin and Timmers 2010; Gabrielsson and Lindström 2010). Yet *portamento* is strongly related to emotional inflections in speech prosody in that the voice often inflects spoken words to modulate it with an emotional meaning (Tao, Kang, and Li 2006; Murray and Arnott 1993). In this paper we shall define *portamento* and argue that it has been present in much singing for important acoustic reasons, and we shall identify some of the factors that may have affected its study. In fact, we identify a paradox of *portamento* — given that the singing voice is inherently more suited to producing *portamento* than stable, fixed pitch (as we shall explain), why is so little know about the emotional effect of *portamento* in singing? *Portamento* is commonplace as an expressive device in all vocal genres we investigated. Several technological, aesthetic and social factors led to the apparent tendency or need to control the amount of *portamento* singers used in some styles of music. This control, plus the belief that *portamento* is not prevalent, plus the difficulties in measuring it account for the '*portamento* deafness' that appears to be present in music psychology research. We conclude by calling for *portamento* in music to be given greater attention in music-emotion research, and consequently we present some hypotheses and discuss methodological issues.

Defining portamento

Portamento refers to a smooth pitch transition, usually from one pitch to another, but could refer to the way a note is attacked or released. The term is used in several ways across the literature. Historically it has origins in the seventeenth century usage *cercar della nota* and *port de voix* (Harris 2014). In an English music dictionary of 1834 we see the entry 'portamento di voce' with its modern meaning (Porter 1834, p. 317). Some researchers equate *portamento* with *glissando* (d'Alessandro, Rosset, and Rossi 1998; Herbert 2010; Lomax 1976)—although, unless convention dictates otherwise, we will consider the second of these as a rapid progression from one pitch to another, like *portamento*, but passing through discrete pitch steps in the direction of the final notated pitch, such as that easily performed by stroking a harp (see also, Burrows 1958). Other terms that have been used to mean *portamento* include scoop (Rycroft 1959; Boyer 1979), Sprechstimme (List 1963), glide (List 1963; Seashore 1938/1967; Schoen 1922), downglides, upglides (Lomax 1976; 1968/2009, p. 56), snaking (Wicks 1989); swoop (a very fast, short version of portamento according to Leech-Wilkinson 2006), pitch inflections (Kauffman 1992; Potter 2006) and sliding (Lomax 1976; Youell 2012; Philip 2004, 'slide' might be considered the nearest English word equivalent to portamento, provided it is clear that we are referring slide in pitch). We shall treat each of these terms as falling under the umbrella of *portamento*.

An acoustic representation of *portamento* is the fundamental frequency (F0) contour that unfolds from a starting time to an end time. Either end can be silent or have a relatively stable F0 value. F0 refers to the fundamental frequency of the note, which is strongly related to the perceived pitch and which in the case of *portamento* necessarily changes with time. The duration of the *portamento* trajectory can vary but to be perceptible would need to have an F0 that changes over a period of at least 20 ms (Rosen and Howell 1987; Hart, Collier, and Cohen 1990), by a ratio greater than the corresponding pitch difference limen (itself a complex condition in terms of acoustic correlates, see Houtsma 1995; de Cheveigné 2010). Further, it would need to fall above the glissando threshold (which is a function of the duration of the glissando, d'Alessandro and Mertens 1995; Hart, Collier, and Cohen 1990). If the pitch transition between two successive notes is continuous but *very* rapid (duration is less than the glissando threshold), it will be classified as *legato* rather than *portamento*.

The acoustic quantification 'F0 contour' and its perceptual counterpart 'intonation contour' or 'melodic contour' are defined by Juslin and Timmers as a "sequence of F0 values. In music, intonation refers to the manner in which the performer approaches and/or maintains the prescribed pitch of notes, in terms of deviation from precise pitch" (Juslin & Timmers, 2010, p. 461). This definition encompasses vibrato, which receives considerably more attention than the varieties of *portamento* (Juslin and Laukka 2003). Other terms that have been used to describe *portamento* effects in the acoustic realm include F0 slope (Goudbeek and Scherer 2010) and microintonations/microvariations (Baroni et al. 1997; for reviews, see Juslin and Persson 2002; Juslin and Timmers 2010). In this paper, melodic contour will be treated as a special case of an F0 sequence, referring to a sequence of stable, discrete pitches occurring over a time scale of seconds, rather than a continuous variation occurring over short time durations (typically 10s to 100s of milliseconds).

Determining whether the transition is legato or *portamento* in musicology is usually done primarily by expert ear (e.g. see Katz 2006). Leech-Wilkinson argues that the swoop (which for Leech-Wilkinson is closely related to *portamento*) is generally shorter than 300 ms. But *portamento* can take as long as a performer wishes, within physical constraints such as the duration of a singer's or wind player's breath and the physical restriction of the instrument such as the length of the fingerboard on a stringed instrument, or the length of a trombone slide. The F0 trajectories observed in *portamento* are typically not periodic (unlike vibrato), and can be monotonic or non-monotonic (i.e., contain one or more local minima/maxima). Given the potential complexities of *portamento*, we will restrict our investigation to those that move directly from one point to another, without overshoot or vibrato (i.e., monotonic).

An established psychoacoustic principle of *portamento* is that perception of it is strongly related to the duration of the slide. d'Alessandro, Rosset & Rossi (1998) found that the larger the F0 range and the longer the time duration between the start and end of the *portamento* transition, the easier the *portamento* was to perceive. These findings were based on a single underlying complex tone with an initial F0 of 220 Hz (corresponding to a pitch of A3). For example, a slide of 1.5 semitones (150 cents) over 50 ms is not perceived as *portamento*, whereas if F0 traverses 6 semitones (600 cents), it is. Similarly if more time is given to the transition, it becomes more likely to be categorised as a *portamento*. In the case of the

1.5 semitone transition, if the duration of the traversal is increased to 200 ms (compared to 50ms), it then becomes perceived as a *portamento*. This pattern outlines the threshold between *legato* and *portamento*, which is better know as the glissando threshold (Hart, Collier, and Cohen 1990).

Consider the excerpt of scat singing by Ella Fitzgerald shown as an F0 plot in Figure 1. The excerpt is from Ella Fitzgerald's c.1958 recording of Blue Skies (Verve 830533-2). The abscissa indicates time elapsed in the track. In the excerpt Fitzgerald is scatting (improvising using nonsense syllables), with a pitch sequence corresponding to the eight syllables shown being E4-C4, E4, C4, G#4-A4-E4-C4, demonstrating a range of transitions, most relevant here being *portamento* (circled F0 segment), *legato* (rectangled F0 segments) and *vibrato* (bracketed F0 segments). The overlaid dotted horizontal line indicates a region of perceptibly stable pitch. Arrows indicate the point at which the syllable begins. Values of gradient and duration for legato boxes are shown adjacent to each in units of semitones per second (ST/s). The inset box shows linear versions of the three identified regions of *portamenti*, with P1 having an F0-slope = 25 ST/s (compared with a glissando threshold, G_{thr}=25), Dur(ation) = 80 ms; P2a having F0-slope = 17 ST/s (G_{thr}=2.6), Dur = 250 ms; P2b having F0-slope = -21 ST/s (G_{thr}=-4), Dur = 200 ms; P3a having F0-slope = 12 ST/s (G_{thr}=4.4), Dur = 84 ms; P3b having F0-slope = 2.2 ST/s (G_{thr}=0.66), Dur = 490 ms. Only P2a, P3a and P3b (i.e. the solid lines in the inset) are considered psychoacoustic *portamenti*, according to the glissando threshold formula of Hart et al. (1990): G_{thr} = 0.16/(T2), where T is the duration of the slide and G_{thr} is in units ST/s. All legato marked regions fall within the glissando threshold, but note that the first legato box (commencing at t = 81.58 s) has a duration of 30 ms, and is close to the threshold. The articulation of the 'd' in 'daa' may help to create a legato percept, in contrast to a diphthong (such as 'dee-aa' instead of 'dee-daa') which may have sounded more portamento-like. As Hart, Collier and Cohen explain:

> In our experience, a drop in amplitude of 10 to 20 dB in several tens of milliseconds, as can often be found in transitions from vowels to consonants, can entirely obscure the presence of changes in F0 of up to half an octave. (Hart, Collier, and Cohen 1990, p. 36)

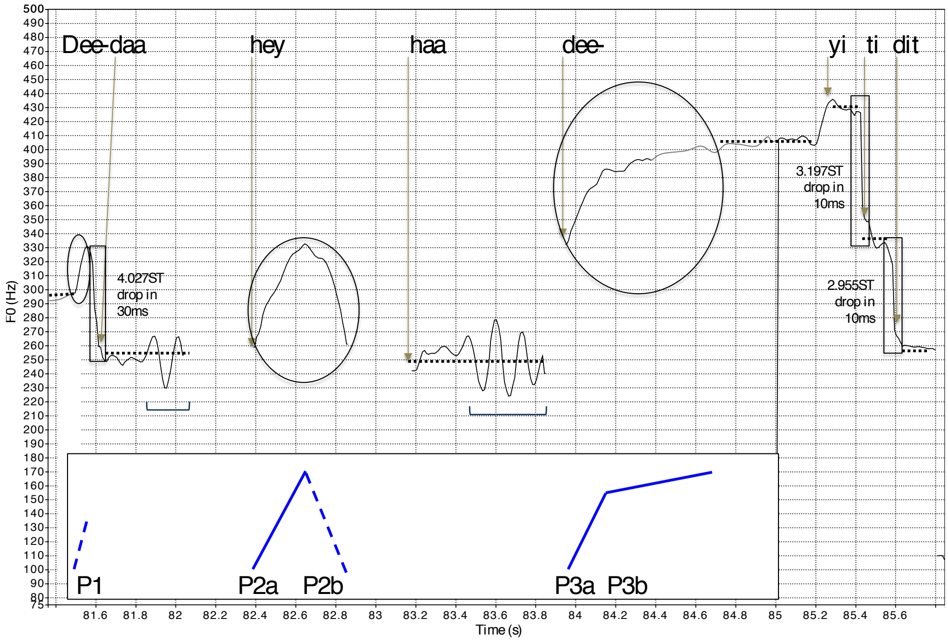

Figure 1. F0 plot from an excerpt of scat singing by Ella Fitzgerald. See text for explanation of annotations.

Emotion and short duration pitch flexibility in speech

The vocal mechanism for producing *portamento* is similar to that for producing inflections in speech (Hoppe et al. 2003). Speech inflections are important for communicating emotional information (Murray and Arnott 1993), whether the language is tonal (Yip 2002) or non-tonal, such as English. Banse and Scherer (1996), for example, summarise the use of downward F0 contour used when expressing 'cold anger' (as distinct from 'hot' anger) and sadness. Trainor and Adams (2000) demonstrated that the pitch contours used for the production of sentences that conveyed love, fear and surprise were different from each other, regardless of whether directed to an infant or to an adult. This provides evidence that emotion is encoded in the changing F0 trajectory of voiced vowels in speech. Mithen (2005) summarises the emotional communication afforded by such inflections through his review of the literature on infant directed speech: "Parents largely do this on an intuitive basis – they do not need to be taught [infant directed speech] – and use music-like utterances and gestures to develop the emotional capacities of the infant prior to facilitating language acquisition" (p. 84).

If we were to accept Mithen's account of the origins of language and speech, *portamento* type effects may well have been an integral part of the protolanguage that was present before contemporary systems of music or language evolved (for a review of protolanguage theories, see Fitch 2006). It would mean that F0 inflections that communicate emotional meaning could be present in both speech and music.

However, the majority of research in music and emotion that relates to speech contour/prosody examines movement of pitch over a longer time, (in the order of seconds, rather than 10s to 100s of milliseconds) and using discrete (fixed, steady), rather than continuous (sliding) pitch as the components of the contour, what is commonly described as melodic contour (for a review of the literature, see Juslin and Laukka 2003).

The research conducted on the relationship between emotion and longer time scale melodic contour is considerably more developed in music psychology. Gabrielsson and Lindstrom (2010) summarised the literature in question by concluding that a rising melodic/pitch contour was related to emotions of fear, surprise, anger, potency and happiness, while falling contours related to sadness, boredom, and pleasantness, with most of the findings drawn from a study by Scherer and Oshinsky (1977) (see also Juslin and Laukka 2003). Thus melodic contour and its emotional correlates are more or less established in the literature, and have undergone scrutiny with respect to their parallels in speech prosody. However, this interest in emotion in melodic contour has not been applied to the short duration, continuous pitch, *portamento* inflections in music performance, despite musicological agreement that they are capable of producing emotional expression, as will be discussed below. We were unable to cite any sources that ruled out the emotional influence of *portamento*, and so in the following section we examine the reasons for this neglect from acoustics, anthropology and history.

Human voice acoustics and portamento

In voiced speech and singing, the vibration of the vocal folds modulates the flow of air from the lungs and this provides the source of sound. The aero-mechanical process that produces the voice is subtle, but we give a brief and simplified version here (for further details, see Titze 1994; Wolfe, Garnier, and Smith 2009; Kob et al. 2011). The positive gauge pressure in the lungs can supply a force below the vocal folds, which can overcome the muscular tension in and across the folds, then separate them and thus allow a pulse of air through the aperture between them (known as the glottis). Two different sorts of forces can then close the folds: one is the elastic forces in the folds, another is the so-called Bernoulli force: the suction due to rapidly moving air through the glottis. The 'appropriate conditions' mentioned above include the values of the tension in the folds, their geometry and the acoustic loads on the folds due to the airways above and below them. Cycles of opening and closing can then repeat, producing a periodic sound wave.

If the geometry, the muscular tension and sub-glottal pressure are all maintained constant, then the vibration cycles are nearly identical (Titze 1973; Wolfe, Garnier, and Smith 2009). This gives rise to a constant F0. (Further, because the vibration is strongly nonlinear, it gives rise to harmonics with frequencies at integral multiples of F0 (Fletcher 1999).)

Increases in sub-glottal pressure increase both the sound level (SL) and F0 (Titze 1973; Wolfe, Garnier, and Smith 2009). Variations in muscular tension also change both these parameters. Pitch depends very largely on F0, and loudness depends on both F0 and the sound level SL. So a change to either the pressure or the tension affects both pitch and loudness. In the normal human voice, F0 and sound level are very strongly correlated. For a singer, a plot of the possible range of SL against F0 is called a Voice Range Profile, and such plots clearly show the positive correlation (Lamesch, Doval, and Castellengo 2012). This is a critical point in our argument. Changes in SL, which normally correspond perceptually to changes in loudness, actually also affects pitch – unless the voice has a significant amount of vocal training and can thus make exact compensation. This is one part of the argument that the voice is better at producing changing pitch than stable pitch.

Thus the voice converts power from a steady source (here the flow of high pressure air from the lungs) into oscillating or acoustic power in a sound wave. Analogous processes can be identified in those artificial instruments that produce sustained, pitched sounds. The analogy is closest for brass (lip reed) instruments, where the forces driving the player's lips have strong similarities with those driving the vocal folds. Reed instruments also use the air pressure provided by the lungs, but with the difference that high-pressure air upstream tends to close the reed. In flutes and the like, pressure variations in a pipe cause an air jet to be deflected alternately to either side of an edge. In bowed string instrument, the steady source of power is the moving bow, and the regular cycle of sticking and slipping between bow and string converts this to a mechanical vibration that drives the bridge and thus the body of the instrument (for more details, see Rossing, Moore, and Wheeler 2002).

However, there is a striking and important difference between the voice and most artificial musical instruments. In artificial instruments, pitch is almost always controlled using a resonator that readily supports vibrations at one or more resonances, each having a rather narrow band of frequencies. For brass or woodwind instruments, the control resonator is the bore of the instrument. For string instruments, it is the string. In these cases, standing waves in the resonator can drive the vibration production mechanism (lips, reed, stick-slip cycle) and thus control F0, typically to rather better than a semitone (e.g. Fletcher and Rossing, 1996). A wind or string instrument player uses his/her fingers to control the effective vibrating length of the bore or string and this configuration allows the production of a small number of notes with relatively stable pitch. Discrete changes in pitch can then be made by changing the positions of the fingers. Further, because of the stabilising effect of the bore or string, louder notes can be produced by blowing harder or by bowing faster with relatively little effect on the pitch.

With the exception of the high soprano range, the human voice does not usually have a resonator to control the pitch. The reason is quantitative: the vocal tract is only about 17 cm long and its *lowest* resonance lies typically in the range 300–800 Hz: from a little above middle C to less than two octaves above (Clark and Yallop, 1990). Although sopranos in their high range tune vocal tract resonances to match pitch (at the cost of phonetic information), altos, tenors and baritones do not use such tuning in a precise or systematic way (Henrich, Smith, and Wolfe 2011; Joliveau, Smith, and Wolfe 2004). Thus, although the vocal tract resonator might be used by sopranos in their high range to assist stable intonation, and perhaps occasionally by other singers, this is not usually the case. Stable pitch singing occurs if all of the relevant variables are held constant. However, fixed F0 with varying subglottal pressure (and thus varying SL) requires precise, compensating control. The flexibility of the vocal tract, in particular as a result of manipulating jaw and tongue position, makes the human voice well suited to the production of formants characteristic of vowels and other phonemes in speech, rather than the fixed pitches that are ubiquitous part of modern Western music. In other words, as we have argued in detail elsewhere (Wolfe and Schubert 2012, 2010), the human voice is suited to producing vowel sounds and other phonemes, where control over formants is essential but where, in non-tonal languages at least, stable, discrete pitch is not so important.[i]

It is worth considering the artificial instruments that don't have a resonator to control the pitch. The musical saw and the theramin are two. *Portamenti* are characteristic on these instruments. Playing fixed steady pitches on them requires care, and playing fast scale passages or arpeggii is extremely difficult. We argue that the only reason why steady pitches and fast passages are not always judged extremely difficult for the voice is in part because performers have literally a whole lifetime of experience in using the voice.

Distinguishing between the vowels ‘u’ and ‘o’ is more important in non-tonal language communication than in distinguishing a steady pitched ‘u’ at 220Hz versus a steady pitched ‘u’ at 300Hz. Many would argue that the distinction between these two comparisons corresponds to the distinction between speech and music, among other factors (List 1963; for a detailed discussion, see Patel 2008). Protolanguages of early hominids may have consisted of what today we would consider both speech-like and music-like qualities, and, if so, prosodic information, communicated through ‘musical’ channels, such as *portamento*, may well have aided in communicating at an emotional level (for further discussion, see Mithen 2009, 2005).

One reason for the split between language and our modern conception of music as based on a fixed gamut of stable pitch combinations (a musical scale) is through the development of artificial instruments, such as the bone flute (Kunej and Turk 2000). Because of the similar geometry, it is expected that the bone flute would have had some playing similarities with today's shakuhachi. Compared with most modern wind instruments, the bone flute probably allowed for relatively easy bending of pitch, over a semitone or so. However, compared with the voice, it allowed relatively easy independence of loudness and pitch and it made very easy the scale or *arpeggii* passages that are rather difficult on the continuous pitch instruments mentioned above, including the voice (Chen, Smith, and Wolfe 2009). The cultural advantages and human desire for mimicry (Garrels 2005) created cultural and cognitive pressure for the human voice to reproduce the sounds of stable pitch artificial musical instruments (Wolfe and Schubert 2010). As a result, music could have begun to take on the sound gamuts that are associated with the now familiar, ubiquitous, fixed, stable-pitch systems. However, as we have argued, the human voice is much better suited to producing continuous pitch, correlated with loudness and has dextrous control of timbre (phonemes), unlike many artificial instruments. Importantly for our argument, one example of this is *portamento*: *Portamento* is both easier for the voice to do than imitating a fixed pitch musical instrument.

Anthropological evidence

We have argued that the human voice is well suited to producing flexible pitch with good control of timbre (such as formants), but without the need or a dedicated mechanism for discrete pitch control. We do not deny that the human voice is capable of producing stable F0. With care – and often after the practice of many *messe di voce* – the voice is capable of stable pitch over large changes in loudness. However, music in a historically modern, Western sense, usually consists of a changing stream of fixed pitches, and often with significant loudness variation during single notes: what singers would call accents, expression and/or phrasing. When the pitch change occurs (what we might refer to as a melody), the biological propensity of the voice is to produce the different pitches *without* great accuracy in terms of evenness in pitch and dynamics, unless some training or at least some exposure to cultural norms have occurred. A majority of infants make pitch flexible vocalisations (Stark 1978) and demonstrate this when learning to sing (e.g., Welch 1994). We argue that vocal music may have started mainly as a pitch flexible system, but through the desire to imitate artificial musical instruments, as discussed above, and a series of further technological developments, many forms of singing continued to adopt features that resembled artificial instrument styles of discrete pitch production.

Evidence for this transfer of fixed pitch instrument characteristics to vocal desires would at least require anthropological accounts of vocal music in the absence, versus in the presence, of fixed pitch artificial instruments. This kind of evidence is difficult to find, but we do know that subcultures and societies without fixed pitch instruments tended to use musical devices that were less reliant on fixed, discrete pitch (List 1963). One of the most important examples, although there are many, is the 'tumbling strain', found for example in tribal central Australian aboriginal communities where songs were often unaccompanied, in the absence of artificial musical instruments such as the didjeridu (which originally was confined to the North of the continent) (Ellis 1965; Moyle 1981). The tumbling strain requires the singer to leap to a high note and gradually cascade down using rattles and slides, ending at a low note (Sachs and Kunst 1965). This device is found in what Sachs refers to as simple cultures, and resembles the act of shouting or crying. Sachs argued that as music becomes more 'evolved', the pitches and rhythms in the tumbling strain become more regular (p. 73). We propose one reason for this is the digitization of pitch: Namely, that as artificial instruments became more prevalent, the interest in imitating them followed. Another is the development of harmony: two or more stable, harmonic sources, whether from the voice or from artificial musical instruments, allow the production of interference and difference tones and thus can lead to harmony and temperaments.

The relative importance of the roles of the voice and other musical instruments in the earliest developments of harmony are lost in prehistory. It seems plausible that the consonance of the octave may have developed as a typical or the 'natural' interval between men's and women's voices in cultures in which both sexes sang and occasionally had sustained notes. (The octave is also, of course, the strongest example of a consonant interval, because the harmonics of the upper note are a subset of those of the lower.) Certainly by the time of the ancient Greeks, the role of string instruments in harmony and temperament was very important (Barbour 1951; West 1992).

As stable pitch systems emerge (as we propose), several factors can be identified which reinforce such systems. Locking the singing voice into a fixed, discrete pitch-gamut appears to have the advantage of allowing harmony (see also Parncutt 2011; Parncutt, Kaiser, and Sapp 2011). Furthermore, the digitisation of pitch allows tunes to be stored and perhaps processed with relatively little information and much reduced sensitivity to noise (Wolfe 2007, 2002). Both are powerful advantages. In principle, both could be consistent with *portamento*: notes with stable pitch could be joined by continuous slides. The stable pitch phases could be used for harmony and for discretisation, and the *portamento* sections interpolated for additional expressive naunce. There are many examples that demonstrate this effect, where a song is performed with *portamento*, but the conventional Western notation, and performances by certain instruments (piano being the most obvious) use discrete pitches only. Consider three examples. (1) The third-last to second-last notes of *Good King Wenceslaus*: These notes, ˆ1-ˆ4, are sometimes sung *portamento*, but written discretely: ˆ1-slide-ˆ4-ˆ1). (2) The first two notes and the next two of the melody of Arlen's *Somewhere Over the Rainbow* (ˆ1-ˆ8 and ^7-^5) usually sung with *portamento* through the diphthong ('whe' in 'somewhere') in the first and 'oh' in 'over' for the second. And (3) on the 'uh' vowel in the word 'summer' of the opening two sung notes of Gershwin's *Summertime* (ˆ5-ˆb3), often sung with a descending *portamento*.

There are alternative explanations for the development of fixed pitch systems. Musical harmony is easier to control in a group setting (more than one performer) if the pitch components of the harmony are stable, and so the development of polyphony may have been a driving force (providing one explanation of why heterophony is relatively rare). But even the presence of harmony may be viewed as an outcome of the original desire to use fixed discrete pitch, rather than being a driving force (for discussion of the further development of fixed pitch, particular with harmony, see Parncutt, Kaiser, and Sapp 2011).

Historical inhibitors

Portamento in singing has been documented throughout much Western Art Music (Strunk 1945; Katsanevaki 2011; Troelsgård 2011; Johnson 1899). For example, in his detailed analysis of vocal motion in ancient music, Johnson discussed the Ancient Greek aesthetic of continuous versus intervallic transitions proposed by Aristoxenus, Ptolemy and Aristides Quintilianus in speech and singing, concluding that "that form of pitch-movement which accompanies the recitation of poetry, as observed by Aristides, consists not only of a musical intonation of the syllables at various degrees of pitch, but also of glides in pitch from degree to degree" (Johnson 1899, p. 49). This quote suggests the absence of *portamento* in singing. But the fact that *portamento* effects are discouraged or criticised can be seen as the evidence of their presence, even if unwelcomed. In support of this interpretation, Katsenevaki quotes

Pseudoplutarch's assertion that some kind of *portamento* is a natural part of the singing voice because "these musicians always soften the *lichanous* and *paranetes*. And they also soften some of the immovable (*estotes*) tones" (cited by Katsanevaki 2011, p. 161, fn 4).

Treatises explicitly identifying *portamento* in singing can be found throughout the common practice period (c.1600-1900), two early examples being Francesco Rognoni's "*Selva de varii passaggi*" (1602) and Christoph Bernhard "*Von der Singe-Kunst, oder Maniera*" (c1649) (Harris 2014). The relatively widespread use of *portamento* has been documented in Western art music in the nineteenth century (Harris 2014; Brown 2004; Kauffman 1992), and the early part of that century is when the term 'portamento' started to become firmly established to describe the effect under investigation {Brown, 2004 #19721, pp. 559-560; See §Definition of *portamento*, above}.

Overuse of *portamento* could be castigated by linking it with 'lower' music forms. Herbert (2010) pointed out that there was an association of crudeness and popularity in the use of *portamento* that had no place in the refined space of the conservatoire. Class-based, societal pressure may have contributed to rid 'good quality' music of the scourge of *portamento* (for further discussion of class relationships influencing popular music, see Wicks 1989; Schubert 2013). This social-political function of *portamento* in distinguishing classes of people is an additional kind of cultural pressure that explains why *portamento* may appear to be become less common, particularly with the rise of the Western middle class. For example, Wicks writes about the attitudes toward folk/popular style singers who 'snake the voice', and the ironic view (from our perspective) that *portamento* is unnatural:

> When Joe S. James, a southern singing school teacher, brought out his edition of the long-lived shape-note hymnal The Sacred Harp in 1911, he stated rather forcefully that his readers would find "but few of the twisted rills and trills of the unnatural snaking of the voice' which he had heard while in the company of those yet untutored in the art of singing by note - or 'regular singing', as it had been called" […]. It was not the first time the 'old way' of singing and its 'rills and trills' had found a critic among champions of some form of the Western European music system, nor would it be the last. Indeed, the history of the quarrel has shown that if there is anything more intransigent than the old way of singing, it is the accompanying opprobrium spread by a musical elite convinced of the superiority of the diatonic system and all that is extrapolated from it. (Wick, 1989, p. 59)

Katz (2006) proposed that the rise of high fidelity sound recording technology was responsible for the aesthetic shift away from the use of *portamento*. He argued that the recorded *portamento* can "sound calculated or contrived when heard repeatedly on record" (p. 225), and that the musicians found that recordings tended to overly exaggerate their *portamento*, particularly without the benefit of visual cues. Leech-Wilkinson (2006) presented an alternate account, arguing that after the turmoil of the two world wars, *portemento* went into decline and 'disappearance' because of deep psychological changes in Western culture. It was *because* of *portamento*'s expressive power, and especially with connections to childhood memories and loving engagement with family that "[p]ortamento is affecting. However much it may irritate or embarrass today" Leech-Wilkinson argues "it is fairly clear that what it tries to do in most cases is to get us to feel more deeply moved than the notes it joins might achieve on their own." (Leech-Wilkinson p. 236). Potter made a similar argument when observing the decline in the use of *portamento* after WWII: "in the bleak cultural mood during and after the war it is not surprising that a performing style that appeared to be extravagant or even decadent would be increasingly disparaged" (Potter 2006, , p. 554)

The advent of high speed printing in the 19th century (McPherson and Gabrielsson 2002) may also have contributed to the suppression of musical *portamento*. Pont (1979) referred to the standardization of musical notation as 'editorial uniformity' which gave the publisher considerable power over the production and dissemination of music. It is easy to see how composer-intended *portamento*-like features, which were difficult to notate accurately, may have disappeared as a result of the transmission of discrete based (discrete pitch notation system) music scores to regions far beyond the places where those performance traditions resided. Harris refers to two singing treatises from the first half of the nineteenth century when she explains:

> J. F. Schubert (*Neue Singe-Schule*, 1804) noted that 'we have no sign in music for this melting of tones into one another', which he … called '*cercar della nota*', and proposed a simple line between notes. Manuel García (*Traité complet de l'art du chant*, 1840–44/R) suggested the slur as a sign for the *port de voix* (or portamento) (Harris 2014).

These two forms of notation—the line and the slur—both found in notation of portamento in the 20th century as well (Brown 1988; Herbert 2010; Milsom 2003a), do not provide explicit information about the rate of change and the duration of the *portamento*, two basic parameters of the device, leaving the execution of even notated *portamenti* to the expressive interpretation of the performer and convention of the day. Janders and Harris remarked:

> A consistent notation for the portamento … was never developed, in part because the practice was so normative that notation would have been redundant. Thus our record of this practice is largely limited to written treatises on singing and descriptions of voices, but early recordings also document its regular use. (Jander and Harris 2014)[ii]

Without an accurate notation, whatever traditions existed in *portamento* performance will be lost, confused, or manipulated and so, without explicit documentation of *portamento*, possible pre-existing traditions may too have been lost (e.g. see Freitas 2002).

The possible restraining effects of music notation and the sound recording are surmised in Taruskin's attack on the authentic performance movement:

> No less than the score, the performance is regarded as a "text" rather than as an activity, and this creates another pressure toward the elimination from it of anything spontaneous or "merely" personal, let alone idiosyncratic. (Taruskin 1995, , p. 61 fn 16)

These apparent straight-jacketing technologies, political circumstances and general criticisms of *portamento* create a discourse of historical decline in the use of *portamento*. Leech-Wilkinson's interpretation suggests another example of the 'general decline' view of changing performance parameters over time, such as the decline of *portamento* in the twentieth century in favour of more prevalent vibrato (Leech-Wilkinson 2009; 2013, p. 46).

But even with these various factors that appear to muffle the use of *portamento* in Western high-art music, we still find its decline or absence is replaced eventually with its rise or rediscovery. Throughout all of the historical periods investigated, disparaging remarks can be found about *portamento* (such as an 1813 article which reported: "it sometimes sounds like the song of dear little cats, which miaow rather than sing" cited by Brown 2004, p. 559, from Wiener allgemeine musikalische Zeitung, Vol. 1, p. 531). Potter (2006) presents a detailed account of the use of *portamento*, and his own analysis of sound recordings of the twentieth century indicated a rise in its use in the early electrical recording era (mid 1920s to mid 1940s), but then a general decline for the rest of the period of analysis (which ended at near the turn of the millennium). However, Potter points out that *portamento* is still used, and its return in violin playing has also been noted recently (Rosand 2014). Clearly there are changes in aesthetic trends that influence the presence, absence and degree of *portamento* used as an expressive device in singing, and so negative and positive views about its use might be expected. Such changing fashions are normal to the psychology of aesthetic appreciation and criticism (Martindale 1999, 1996, 1990; Simonton 1998b, 1998a).

Our interpretation[iii] of the evidence is that for Western high-art music the use of *portamento* in singing went in and out of fashion. For example, although much of the nineteenth century may be regarded as a period of liberal use of *portamento* in high-art music, Freitas identified a source that hints at an undulation in taste for *portamento* in the middle of the nineteenth century through a secondary source referring to "an article in the *Gazzetta Musicale*, written by Alberto Mazzucato in 1842 …, that claims the practice of portamento had by then almost died out" (Freitas 2002, , p. 244, fn. 89). The waxing and waning, rather than the systematic decline, of many artistic activities may be better explained in such a way. This non-linear pattern in music performance styles has its own supporters (e.g. Fabian 2003, p. 246, who refers to the pattern of changes in taste as a "Hegelian spiral").

Thus the anthropological and historical evidence suggests that social and technological pressures inhibited flexible pitch delivery in music, but since the voice is so well designed to produce it and since it is able to convey important affective information, flexible pitch delivery, if ever stamped out, will find ways to reappear, usually under the guise of expressive nuance and ever-undulating aesthetic taste. Furthermore, as stable pitch instrument accompanying ensemble sizes grew, particularly in opera, the singer could utilise features of the voice that give it an advantage over the increasingly louder orchestral backing, by using vibrato, the singers formant and, here, we claim, possibly even *portamento*. Regardless, *portamento* is more alive in the expressive tool kit of the singer (and other continuous pitch instruments) than it is in the realisation of notated *portamento*. The human vocal apparatus is well designed to use it as an expressive device but, to date, Western music notation is not apt to document it sufficiently.

Our conclusion is that *portamento*, even if repressed in some musical styles at certain times, does not diminish its importance as an expressive device. At the very least, if it is, we should find out why. Put together with *portamento's* likelihood to contribute to emotional meaning, the examination of performance parameters in music should be expanded to include *portamento* to see what expressive outputs can be systematically identified. We therefore briefly describe four issues for the experimental investigation of emotion in *portamento*: (1) the kind of results that might be expected/tested, (2) problems with transferring findings from speech to hypotheses in music, (3) issues regarding the coding of *portamento* and finally (4) the selection of stimuli and their relative impact in comparison to other musical features.

Hypothesis generation

The link between musical *portamento* and the human, speaking voice has been acknowledged for some time[iv], and therefore provides an opportunity for generating hypotheses. Potter claims that:

> Gliding from note to note has rhetorical implications for tempo: the singer can control the pace of the phrase by sliding, and has the possibility of introducing para-linguistic tropes such as sighing, sobbing, and other effects designed to manage the rhetorical communication of emotion. (Potter 2006, , p. 549)

But while the music history literature cited above frequently acknowledge the emotional capacity of *portamento*, it rarely presents or refutes the *specific* emotions that different kinds of *portamenti* produce, meaning that we are somewhat limited in generating hypotheses for testing. Milson's assembly of late nineteenth century expressive devices suggest that *portamento* was associated with passionate character (p, 84), but could also be used to generally heighten emotion:

> Recordings by Adelina Patti (1843-1919) suggest that, for this singer ..., mood and character suitability provide the chief justification for the use of portamento. The relatively restrained mood context of '*Batti. batti*', from *Don Giovanni*, creates a … cerebral context … and, accordingly, the device is used more sparingly. […] Moreover, Patti's slides are usually fast and light, and where they are not, they serve either to heighten the emotion (as at the end) or simply as an expression of appropriate '*joie de vivre*'. (Milsom 2003b, , p. 89)

One reasonably specific writing that links different kinds of *portamenti* to different emotions is Garcia's nineteenth century treatise on singing. It provides explicit observations of the emotions singers may expect to produce with different kinds of *portamenti*

> Slurring [portamento] is a method—sometimes energetic, sometimes graceful,—in colouring a melody; when applied to the expression of forcible sentiments, it should be strong, full, and rapid … When used in tender and graceful passages, it must be slower and softer … Yet it may be observed, that a slur will always be well placed, whenever, in passionate passages, the voice drags itself on under the influence of a strong or tender sentiment (Garcia 1924/1872, p. 53).

Monelle (2012) in citing Baillot reported that *portamento* produced tender expression, provided it was not overused (p. 66)[v]. Such examples indicate that *portamento* may be able to communicate a variety of emotions and in particular the expression of feelings such as sadness and passion, and that there may be an optimal number of times it could be used to capitalise on these effects. Furthermore, Garcia shows sensitivity to its time dependent characteristics.

As discussed above (§Emotion and short duration pitch flexibility in speech), more specific predictions on the effects of *portamento* can be drawn from the literature on emotional prosody in speech. Papousek's (Papoušek 1994; Papoušek, Papoušek, and Symmes 1991) examination of infant directed speech identified different kinds of F0 contours depending on the emotional arousal being conveyed by the parent. High arousal emotions are conveyed by steep slope – for example, falling sharply to communicate a warning and prohibition, and rising sharply to attract attention, while low arousal (soothing) emotion elicitation has F0 contour characterised by a flat slope. These contours last between 100 and 800 ms, even though they are referred to as melodic contours, and therefore are in the temporal order of magnitude in which *portamenti* can be expected.

Data pertinent to the time varying aspects of *portamento* have been assembled in Table 1, based on literature on emotional prosody by (1) Scherer, Johnstone & Klasmeyer (2002)—a fairly broad review, (2) Ververidis and Kotropoulos (2006)—a review with stringent inclusion criteria, and (3) Pierre-Yves (2003)—a study that acknowledges the importance of short duration time-varying parameters on emotion in speech. In addition, the table includes (4) relevant findings from a review of expressive devices used in music performance by Juslin and Timmers (2010) and (5) the results reported in the review of Gabrielesson and Lindström (2010) with regard to melody contour (i.e. movement of discrete pitch melodies over longer time spans than that of *portamento*). Finally, (6) the introspections and observations in Garcia's treatise are included.

Interpreting these findings in terms of *portamento* in singing, we would predict that high arousal emotions (stress, happiness, anger), shown in the first three data columns of Table 1, will be reflected by *portamento* that produces a wide leap or fall in pitch, and does so rapidly (large gradient). High arousal emotions will be supported by a fast, sudden drop in pitch, although these results are not all consistent. The sadness emotion column in Table 1 has the most consistent results across the literature, and points to *portamento* at the end of the note, which is characterised by a falling or fairly flat contour. Garcia implies that the duration of such F0 profiles would be longer than the high arousal counterparts, but this may be in part because low arousal music is generally performed at a slower tempo (Gabrielsson and Lindström 2010; Schubert 2004). The other columns of emotions shown in Table 1 have inconsistent results or too few data points, but a clear, consistent pattern is evident for happiness and sadness, partly because they are the most frequently studied.

Table 1. Comparison of speech and music literature on time varying short duration acoustic and music emotion eliciting parameters

		Emotion						
Study	Acoustic Parameter	Happiness/ Elation/ Joy	Arousal/Stress	Anger/Rage	Fear/Panic	Sadness	Boredom	Comfort/ Pleasantness
Scherer et al. (2002)*	F0 range	>	>	>	<>	<	<=	-
	Gradient of F0 rising and falling	>	>	>	<>	<	<=	-
	F0 final fall: range and gradient	>	>	>	<>	<	<=	-
Ververidis & Kotropoulos (2006)*	Range	>	-	-	-	<	-	-
	Contour	!	-	-	"	"	-	-
	Transmission duration	<	-	<	<	>	-	-
Pierre-Yves (2003)	Contour	!	-	"	-	"	-	!
Juslin & Timmers (2010)*	Microintonation	!	-	-	Sharp	Flat	-	-
Gabrielsson & Lindstrom (2010)*	Pitch contour	!	-	!	!	"	"	"
Garcia (1924/1872)	Based on descriptions from teaching manual	Rapid (for 'graceful')	-	-	-	Slow and soft (for 'tender')	-	-

Note.

- * literature review
- \> less or smaller than for a neutral stimulus
- < greater or larger than for a neutral stimulus
- <> different to a neutral stimulus
- <= same or greater/larger than for a neutral stimulus
- = same as for a neutral stimulus
- ! rising/inclining (positive F0 slope)
- " falling/declining (negative F0 slope)
- \- no relevant data reported

Problems with translating findings from speech to music

It should be noted that when translating findings from speech research into hypotheses in music research, several restrictions must be noted. Tempo (speed) and dynamics (loudness) are variables found in both music and speech. In music, however, pitch is the focal drive for F0 and controlled by the demands of rhythm, and so the window for curving F0 is often smaller and more restricted than it is for speech. Furthermore, there is an ambiguity in the 'final fall', which for speech is usually the end of a sentence and has semantic implications (e.g. a question will generally have a final rise), but for music, where denotative meaning is not the norm, the reasoning for the placement of the *portamento* might be an exclusively aesthetic one. This is why a simple rule is unlikely to account for much variance in emotion through *portamento* – for example, more frequent use of sadness associated *portamento* is unlikely to produce increasing sadness, consistent with the declaration of many singing treatises to not overuse *portamento* (see fn v). Even within spoken language, the effects of individual parameters are not always independent and reducible, and so caution is required in the design of the study. Different ways of delivering the intonation of a sentence can have drastic implications on semantics, and do not always communicate emotions. As Scherer *et al* warn:

> final-rise versus final-fall patterns of F0 in themselves do not carry emotional meanings, but they are linked to sentence modes such as question versus non-question. However, it can be shown that context, such as type of sentence, affects interpretation. While a falling intonation contour is judged as neutral in a WH-question, it is judged as aggressive or challenging in a yes/no question. (Scherer, Johnstone, and Klasmeyer 2002, p. 447)

Coding Portamento

Attempting to identify hypotheses for relating the time-varying aspects of *portamenti* to emotions is quite scant in music psychology research. As Desain and Honing lament "what happens during and in between the notes is sometimes more relevant than the realization of the note onsets themselves … but this issue is not often addressed in music psychology" (Desain and Honing 1995, p. 393; see also Gabrielsson and Lindström 2010, p. 392; Schubert 2001). In other words, one of the reasons for the lack of interest in measuring the influence of *portamento* on emotion in music is because it is a percept that depends upon temporally unfolding characteristics, and therefore cannot, like many other variables, be meaningfully summarized with a mean value. In addition to the matter of F0 contour being used to refer to short duration inflections as distinct from longer time pitch trajectories of melodies, Bänziger and Scherer (2005) note that of the 104 studies reviewed by Juslin and Laukka (2003):

> …only 25 studies included a description of F0 contours or an attempt to influence emotional attributions through the systematic manipulation of F0 contours. When descriptions of F0 contours are provided they often come down to global 'rising' or 'falling' of the overall contour shape. (Bänziger and Scherer 2005, p. 254)

Simplified approaches are also found in musicological investigations. There, introspections, 'ears' and a score, and in more recent times visualisation software (Cook and Leech-Wilkinson 2009) are the primary research tools (in particular, Brown 1988; Katz 2006; Kauffman 1992; Leech-Wilkinson 2006; Potter 2006).[vi] For instance, Potter describes the analysis method of one of his studies:

> I listened to ninety-seven recorded performances from the beginning of the twentieth century to 2002, noting where each example occurred, distinguishing between ascending and descending forms, and recording the total number of identifiable portamenti in each one. (Potter 2006, p. 540)

Experimental (music) psychology researchers tend to reduce a variable to a simple, reliably codable, preferably objective set of levels, and then over the course of additional studies and theory building, ramp up its complexity to allow for more sophisticated enquiries (Creswell 2003; Maxwell and Delaney 2004). For example, in a cross cultural study of children's response to music, Adachi, Trehub and Abe coded *portamento* as present or absent, and concluded that "Canadian and Japanese children interpreted the renditions as sadder in the presence of portamento." (Adachi, Trehub, and Abe 2004, p. 330). As another example, Dry and Gabrielsson (1997) asked a four-piece rock 'n roll combo (which included a vocalist) to identify the presence or absence of *portamento* (they called it *glissando*). Its presence was associated with the expression of sadness and tenderness, and some with the expression of happiness, but it was infrequent when expressing anger. That is, the simplified coding of *portamento* has led to meaningful findings.

But such a simplified characterization, while convenient, does not code the duration or gradient of the F0 projection, and therefore may explain why some results are inconsistent (Table 1). As a minimum, two parameters (degrees of freedom) are required to provide the most basic time dependent characterization of *portamento*, such as the F0 slope (usually in units of semitones per second) and duration of climb or fall (in seconds or milliseconds). For slower *portamenti*, the shape of the F0 trajectory may become a further issue, because the slope may commence gently and then increase suddenly, or vice-versa, or remain (approximately) linear. The threshold between what constitutes a slow and a fast *portamento* also needs to be determined. But slow *portamenti* could be broken up into separate, linear segments (for an example, see Figure 1).

The time-varying aspects of F0 are attracting increasing interest in speech analysis (Sethu, Ambikairajah, and Epps 2013), and have been shown to be important determinants of emotion in speech (Pierre-Yves 2003, p. 160). However, beyond this minimum, the number of parameters related to *portamento* are numerous. In musicological writing these parameters have been described as follows:

> The amount of time taken by the portamento, the rate of the movement between notes, the accompanying dynamic shading and, contrary to the admonitions of the pedagogues, the placement of the syllable together provide a myriad of ways to approach and quit a note. (Freitas 2002, p. 244)

The nature of the parameters may be worth considering, too. Currently, the F0-contour is a fairly entrenched variable used for identifying *portamento* because of its ease of access via digital audio analysis software. However, other time based approaches are emerging which may have advantages in terms of correlating with emotion, even if the equipment for collecting such data is not as widespread. For example, the time-unfolding measures of physical glottal parameters—directly related to the physics of generating F0 inflections—have revealed good predictive utility of emotional nuance (Cummings and Clements 1995; Sethu, Ambikairajah, and Epps 2013).

Effect size and stimuli

Another reason that *portamento* has not received more attention is because its impact, in comparison with other musical features upon emotion, may be small, even if one were statistically observable, and just as Juslin and Laukka conclude in their review of vibrato and emotion literature (Juslin and Laukka 2003, p. 797), *portamento* might also not *systematically* communicate particular emotions. This issue needs to be considered in designs where *portamento* is controlled by (1) generating a contrived, synthetic signal, (2) manipulating existing recordings through digital editing techniques, or (3) adjusting, as instructed, a highly skilled performer in two or more versions of a piece that is otherwise performed identically (e.g. one version using some *portamento*, and another with discrete pitch and legato transitions only). All of these approaches are likely to be needed before we can conclude that *portamento* has an effect on emotional response independently of other musical features, or interacting with other features. Consider again the example of Fitzgerald's scat singing (Figure 1). The excerpt conveys a fun, joyous outpouring. But are other features (major harmony, fast tempo, relatively high pitch) sufficient to swamp any mood altering effects of the *portamenti*? Under such a circumstance would replacing *portamento* episodes with only discrete steps (as already present in the 'vi-ti-dit' utterance at the end of the excerpt) make a difference to the mood rating? Furthermore, should the long slide (P3) in Figure 1 not have been performed faster to be consistent with the data in Table 1?

Conclusions

This paper has argued that although the human voice has a natural proclivity to use flexible pitch in music because of its physiology and consequent acoustic properties, several musical and historical developments and psychological pressures have led to its suppression in some musical styles. The first, anthropologically important pressure, we speculate, is the discovery and development of the stable pitch, artificial musical instrument. Having acoustic properties almost complementary to those of the human voice, artificial musical instruments provided a model for mimicking, a medium for development of harmony and therefore an attenuation of *portamento* in music, as distinct from speech. We identified four, non-independent, further reasons for the mitigation of *portamento* in music: (1) fluctuations in aesthetic tastes, (2) development of widespread music notation and printing technology, (3) the widespread availability of sound recordings and (4) social/political factors, such as elitism and response to human catastrophe.

However, we also found evidence that flexible pitch survived and at times flourished despite these pressures. *Portamento* appears in many vocal forms found in current popular music, including Gospel, Soul, Blues and Jazz (Wicks 1989; Reilly 1973; Yurchenco 1995; Legg 2010). It was probably always prevalent in folk music (consider Bob Dylan, for example, whose approximate pitching is part of his folk style). We also found evidence that it existed and still exists in Western high-art music despite some arguments of its decline.

The flexible intonations used in natural language have the capacity to communicate emotional nuances, and this 'emotional prosody' has received considerable attention in the research literature. The lack of investigation of the pitch flexibility in music psychology we therefore attribute to the factors of the *apparent* decline of *portamento*, and the difficulty in examining transient parameters of music expression. Studies cited examining the emotional impact of *portamento* are restricted to reduced representations such as 'absent' or 'present', and we call for a more sophisticated approach, which reports as a minimum the duration and slope of the F0 transition that constitutes *portamento*.

This paper has aimed to identify *portamento* as a relevant feature in the vocal communication of emotion that has been neglected in music-psychology. The research agenda this paper proposes is that *portamento* receive serious attention in experimental music psychology to allow a more complete understanding of emotion in music. By applying findings about emotion prosody in speech and from singing treatises, we predict that the portrayal of sadness in music will be enhanced with occasional downward inflections of relatively long duration, particularly at the end of a note, and that the portrayal of happiness will be enhanced by upward inflections over a short time duration. Furthermore, we expect that the number of *portamenti* used will not be in direct proportion to the addition of the emotion that the single instance predicts (e.g. adding more sad-inducing *portamenti* will not lead to a judgement of the piece being commensurately more sad). That is, the number of occurrences and nature of a particular kind of *portamento* is likely to influence the emotion it supports in comparison to the use of the same kind of *portamento* when used in isolation. We also propose that quantification of time durations for *portamento* can be based, as an initial simplifying step, on the glissando threshold. With only two emotions predicted that have any consistency in the literature (happiness and sadness), and so many parameters that describe *portamento*, there is clearly a lot to be done on this topic.

Acknowledgement

The authors are grateful to the two anonymous reviewers of an earlier version of this publication. This research was supported by the Australian Research Council.

References

Adachi, Mayumi, Sandra E Trehub, and Jun Ichi Abe. 2004. "Perceiving emotion in children's songs across age and culture1." *Japanese Psychological Research* no. 46 (4):322-336.

Bänziger, Tanja, and Klaus R Scherer. 2005. "The role of intonation in emotional expressions." *Speech communication* no. 46 (3):252-267.

Barbour, J Murray. 1951. *Tuning and temperament: A historical survey*. NY: Dover.

Baroni, M, R Caterina, F Regazzi, and G Zanarini. 1997. Emotional aspects of singing voice. Paper read at Proceedings of the III Triennial ESCOM Conference. Uppsala: University of Uppsala, at Uppsala, Sweden.

Boyer, Horace Clarence. 1979. "Contemporary gospel music." *The Black Perspective in Music* no. 7 (1):5-58.

Brown, Clive. 1988. "Bowing styles, vibrato and portamento in nineteenth-century violin playing." *Journal of the Royal Musical Association* no. 113 (1):97-128.

———. 2004. *Classical and Romantic performing practice 1750-1900*. Oxford: Oxford University Press.

Burrows, Edwin Grant. 1958. "Music on Ifaluk Atoll in the Caroline Islands." *Ethnomusicology* no. 2 (1):9-22.

Chen, Jer-Ming, John Smith, and Joe Wolfe. 2009. "Pitch bending and glissandi on the clarinet: roles of the vocal tract and partial tone hole closure." *The Journal of the Acoustical Society of America* no. 126 (3):1511-1520.

Cook, Nicholas, and Daniel Leech-Wilkinson. 2014. *A musicologist's guide to Sonic Visualiser*. Centre for the History and Analysis of Recorded Music 2009 [cited 14 February 2014]. Available from http://www.charm.kcl.ac.uk/analysing/p9_1.html.

Creswell, John W. 2003. *Research design: Qualitative, quantitative, and mixed methods approaches*. CA: Sage.

Cummings, Kathleen E, and Mark A Clements. 1995. "Analysis of the glottal excitation of emotionally styled and stressed speech." *The Journal of the Acoustical Society of America* no. 98 (1):88-98.

d'Alessandro, Christophe, and Piet Mertens. 1995. "Automatic pitch contour stylization using a model of tonal perception." *Computer Speech & Language* no. 9 (3):257-288.

d'Alessandro, Christophe, Sophie Rosset, and Jean-Pierre Rossi. 1998. "The pitch of short-duration fundamental frequency glissandos." *The Journal of the Acoustical Society of America* no. 104 (4):2339-2348.

de Cheveigné, Alain. 2010. "Pitch perception." In *The oxford handbook of auditory science: Hearing*, edited by Christopher J. Plack, 71-104.

Desain, Peter, and Henkjan Honing. 1995. Towards algorithmic descriptions of continuos modulations of musical parameters. Paper read at Proceedings of the 1995 International Computer Music Conference, at San Francisco.

Dry, Andrew, and Alf Gabrielsson. 1997. Emotional expression in guitar band performance. Paper read at Third Trienniel ESCOM Conference, 7- 12 June, 1997, at Uppsala, Sweden.

Ellis, Catherine J. 1965. "Pre-instrumental scales." *Ethnomusicology* no. 9 (2):126-137.

Fabian, Dorottya. 2003. *Bach Performance Practice 1945-1975: A Comprehensive Review of Sound Recordings and Literature*. Aldershot: Ashgate.

Fitch, W Tecumseh. 2006. "On the biology and evolution of music." *Music Perception* no. 24 (1):85-88.

Fletcher, Neville H. 1999. "The nonlinear physics of musical instruments." *Reports on Progress in Physics* no. 62 (5):723-764.

Freitas, Roger. 2002. "Towards a Verdian ideal of singing: Emancipation from modern orthodoxy." *Journal of the Royal Musical Association* no. 127 (2):226-257.

Gabrielsson, Alf, and E. Lindström. 2010. "The role of structure in the musical expression of emotions." In *Handbook of Music and Emotion: Theory, Research, Applications*, edited by Patrik N. Juslin and John Sloboda, 367-399. Oxford: OUP.

Garcia, Manuel. 1924/1872. *Garcia's treatise on the art of singing: A compendious method of instruction, with examples and exercises for the cultivation of the voice*. London: Leonard & Co.

Garrels, Scott R. 2005. "Imitation, mirror neurons, and mimetic desire: Convergence between the mimetic theory of René Girard and empirical research on imitation." *Contagion: Journal of Violence, Mimesis, and Culture* no. 12 (1):47-86.

Goudbeek, Martijn, and Klaus R. Scherer. 2010. "Beyond arousal: Valence and potency/control cues in the vocal expression of emotion." *The Journal of the Acoustical Society of America* no. 128:1322.

Harris, Ellen T. 2014. "Portamento." In *Grove Music Online*. *Oxford Music Online*. retrieved from http://www.oxfordmusiconline.com/subscriber/article/grove/music/40990: Oxford University Press.

Hart, Johan 't, René Collier, and Antonie Cohen. 1990. *A perceptual study of intonation: An experimental-phonetic approach to speech melody*. Cambridge: Cambridge University Press.

Hast, Harry Gregory. 1929. "Vocal and unvocal." *Music and Letters* no. 10 (3):250-253.

Hast, Harry Gregory. 1925. *The Singer's Art: Letters from a singing master*. London: Methuen.

Henrich, Nathalie, John Smith, and Joe Wolfe. 2011. "Vocal tract resonances in singing: Strategies used by sopranos, altos, tenors, and baritones." *The Journal of the Acoustical Society of America* no. 129 (2):1024-1035.

Herbert, Trevor. 2010. "Trombone Glissando: A Case Study in Continuity and Change in Brass Instrument Performance Idioms." *Historic Brass Society Journal* no. 22:1-18.

Hoppe, Ulrich, Frank Rosanowski, Michael Döllinger, Jörg Lohscheller, Maria Schuster, and Ulrich Eysholdt. 2003. "Glissando: laryngeal motorics and acoustics." *Journal of Voice* no. 17 (3):370-376.

Houtsma, Adrianus JM. 1995. "Pitch perception." In *Hearing*, edited by Brian C. J. Moore. San Diego: Academic Press

Jander, Owen, and Ellen T. Harris. 2014. Singing, §6. Performing practice. In *Grove Music Online*. *Oxford Music Online*. Oxford University Press.

Johnson, Charles W. L. 1899. "The Motion of the Voice, ἡ τῆς φωνῆς# x03BA; ίνησις, in the Theory of Ancient Music." *Transactions and Proceedings of the American Philological Association* no. 30:42-55.

Joliveau, Elodie, John Smith, and Joe Wolfe. 2004. "Acoustics: tuning of vocal tract resonance by sopranos." *Nature* no. 427 (6970):116-116.

Juslin, Patrik N, and Roland S Persson. 2002. "Emotional communication." In *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning: Creative Strategies for Teaching and Learning*, edited by Richard Parncutt and Gary E. McPherson, 219-236. OUP: Oxford.

Juslin, Patrik N, and Renee Timmers. 2010. "Expression and communication of emotion in music performance." In *Handbook of Music and Emotion: Theory, Research, Applications*, edited by P. N. Juslin and J. A. Sloboda, 453-489. Oxford: OUP.

Juslin, Patrik N., and Petri Laukka. 2003. "Communication of emotions in vocal expression and music performance: Different channels, same code?" *Psychological Bulletin* no. 129 (5):770-814.

Juslin, Patrik N., and John A. Sloboda. 2010. *Handbook of Music and Emotion: Theory, research, applications*. Oxford: OUP.

Katsanevaki, Athena. 2011. "Chromaticism: A theoretical construction or a practical transformation?" *Muzikologija* (11):159-180.
Katz, Mark. 2006. "Portamento and the Phonograph Effect." *Journal of Musicological Research* no. 25 (3-4):211-232.
Kauffman, Deborah. 1992. "Portamento in Romantic Opera." *Performance Practice Review* no. 5 (2):139-158.
Kob, Malte, Nathalie Henrich, Hanspeter Herzel, David Howard, Isao Tokuda, and Joe Wolfe. 2011. "Analysing and understanding the singing voice: Recent progress and open questions." *Current Bioinformatics* no. 6 (3):362-374.
Kunej, Drago, and Ivan Turk. 2000. "New perspectives on the beginnings of music: Archaeological and musicological analysis of a Middle Paleolithic bone 'flute'." In *The origins of music*, edited by Nils L. Wallin, Björn Merker and Steven Brown, 235-268. MIT Press.
Lamesch, Sylvain, Boris Doval, and Michèle Castellengo. 2012. "Toward a more informative voice range profile: The role of laryngeal vibratory mechanisms on vowels dynamic range." *Journal of Voice* no. 26 (5):672. e9-672. e18.
Leech-Wilkinson, Daniel. 2006. "Portamento and Musical Meaning." *Journal of Musicological Research* no. 25 (3-4):233-261.
———. 2009. "Recordings and histories of performance style." In *The Cambridge companion to recorded music*, edited by N. Cook, E. Clarke, D. Leech-Wilkinson and J. Rink, 246-262. Cambridge: Cambridge University Press.
———. 2013. "The emotional power of musical performance." In *The Emotional Power of Music: Multidisciplinary perspectives on musical arousal, expression, and social control*, edited by Tom Cochrane, Bernardino Fantini and Klaus R. Scherer, 41-54. Oxford: Oxford University Press.
Legg, Andrew. 2010. "A taxonomy of musical gesture in African American gospel music." *Popular Music* no. 29 (1):103-129.
List, George. 1963. "The boundaries of speech and song." *Ethnomusicology* no. 7 (1):1-16.
Lomax, Alan. 1968/2009. *Folk song style and culture*. NJ: Transaction Publishers.
———. 1976. *Cantometrics: an approach to the anthropology of music*. Berkeley: University of California Extension Media Center.
Maher, Robert C. 2008. "Control of synthesized vibrato during portamento musical pitch transitions." *Journal of the Audio Engineering Society* no. 56 (1/2):18-27.
Martindale, Colin. 1990. *The clockwork muse: The predictability of artistic change*. New York: Basic Books.
———. 1996. "Psychological determinants of artistic change." *International Journal of Psychology* no. 31 (3-4):2114-2114.
———. 1999. "Peak shift, prototypicality and aesthetic preference." *Journal of Consciousness Studies* no. 6 (6-7):52-54.
Maxwell, Scott E, and Harold D Delaney. 2004. *Designing experiments and analyzing data: A model comparison perspective*. NY: Psychology Press.
McPherson, Gary E, and Alf Gabrielsson. 2002. "From sound to sign." In *The science and psychology of music performance: Creative strategies for teaching and learning*, 99-116. Oxford: OUP.
Milsom, David. 2003a. *Theory and practice in late nineteenth-century violin performance: an examination of style in performance, 1850-1900*. Aldershot: Ashgate.
———. 2003b. *Theory and Practice in Late Nineteenth-Century Violin Performance: An Examination of Style in Performance, 1850-1900*. Aldershot: Ashgate.
Mithen, Steven. 2005. *The Singing Neanderthals: The Origins of Music, Language, Mind and Body*. London: Weidenfeld & Nicolson.
———. 2009. "Holistic communication and the coevolution of language and music: resurrecting an old idea." In *The prehistory of language*, 58.

Monelle, Raymond. 2012. "The Orchestral String Portamento as Expressive Topic." *Journal of Musicological Research* no. 31 (2-3):138-146.

Moyle, Alice M. 1981. "The Australian didjeridu: A late musical intrusion." *World Archaeology* no. 12 (3):321-331.

Murray, Iain R, and John L Arnott. 1993. "Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion." *The Journal of the Acoustical Society of America* no. 93:1097.

Papoušek, Mechthild. 1994. "Melodies in caregivers' speech: A species specific guidance towards language." *Early Development and Parenting* no. 3 (1):5-17.

Papoušek, Mechthild, Hanuš Papoušek, and David Symmes. 1991. "The meanings of melodies in motherese in tone and stress languages." *Infant behavior and development* no. 14 (4):415-440.

Parncutt, Richard. 2011. "The Tonic as Triad: Key Profiles as Pitch Salience Profiles of Tonic Triads." *Music Perception* no. 28 (4):333-366.

Parncutt, Richard, Fabio Kaiser, and Craig Sapp. 2011. "Historical development of tonal syntax: Counting pitch-class sets in 13th-16th century polyphonic vocal music." In *Mathematics and computation in music*, 366-369. Springer.

Patel, Aniruddh D. 2008. *Music, lanugage, and the brain*. Oxford: Oxford University Press.

Philip, Robert. 2004. *Performing music in the age of recording*. New Haven, CT: Yale University Press.

Pierre-Yves, Oudeyer. 2003. "The production and recognition of emotions in speech: features and algorithms." *International Journal of Human-Computer Studies* no. 59 (1):157-183.

Pont, Graham. 1979. "A revolution in the science and practice of music." *Musicology Australia* no. 5 (1):1-66.

Porter, William Smith. 1834. *The Musical Cyclopedia: On The Principles of Music Considered as a Science and an Art: Embracing a Complete Musical Dictionary, and the Outlines of a Musical Grammar, and of the Theory of Sounds and Laws of Harmony, with Directions for the Practice of Vocal and Instrumental Music, and a Description of Musical Instruments*. Boston: James Loring.

Potter, J. 2006. "Beggar at the Door: the rise and fall of portamento in singing." *Music and Letters* no. 87 (4):523-550.

Reilly, April. 1973. "The Impact of Technology on Rhythm'n'Blues." *The Black Perspective in Music*:136-146.

Rosand, Aaron. *Aaron Rosand on portamento: The American violinist examines how the great 20th-century violinists used the 'gliding technique' to bring personality and individuality to their interpretations*, Wednesday, 12 February 2014. Available from http://thestrad.com/latest/blogs/aaron-rosand-on-portamento.

Rosen, Stuart, and Peter Howell. 1987. "Is there a natural sensitivity at 20 ms in relative tone-onset-time continua? A reanalysis of Hirsh's (1959) data." In *The psychophysics of speech perception*, edited by Schouten M. E. H., 199-209. Boston:: Martinus Nijhoff.

Rossing, Thomas D., Richard F. Moore, and Paul A. Wheeler. 2002. *The Science of Sound*. 3rd ed. San Francisco: Addison-Wesley.

Rycroft, David. 1959. "African music in Johannesburg: African and non-African features." *Journal of the International Folk Music Council* no. 11:25-30.

Sachs, C., and J. Kunst. 1965. *The wellsprings of music*. New York: McGraw-Hill.

Scherer, Klaus R., Tom Johnstone, and Gundrun Klasmeyer. 2002. "Vocal expression of emotion." In *Handbook of Affective Sciences*, edited by Davidson R.J., Goldsmith H. and Scherer Klaus R., 433-456. New York: Oxford University Press.

Scherer, Klaus R., and James S. Oshinsky. 1977. "Cue utilization in emotion attribution from auditory stimuli." *Motivation and Emotion* no. 1 (4):331-346.

Schoen, Max. 1922. "An experimental study of the pitch factor in artistic singing." *Psychological Monographs* no. 31 (1):230-259.

Schubert, Emery. 2004. "Modeling perceived emotion with continuous musical features." *Music Perception* no. 21:561-585.

Schubert, Emery. 2001. "Continuous measurement of self-report emotional response to music." In *Music and emotion: Theory and research*, edited by Patrik N. Juslin and John A. Sloboda, 393-414. Oxford: Oxford University Press.

———. 2013. "Emotion in popular music: A psychological perspective." *Volume! La revue des musiques populaires - The journal of popular music studies* no. 10 (1):265-285.

Seashore, C. E. . 1938/1967. *Psychology of music*. New York: McGraw Hill.

Sethu, Vidhyasaharan, Eliathamby Ambikairajah, and Julien Epps. 2013. "On the use of speech parameter contours for emotion recognition." *EURASIP Journal on Audio, Speech, and Music Processing* no. 2013 (1):1-14.

Simonton, D. K. 1998a. "Fickle fashion versus immortal fame: Transhistorical assessments of creative products in the opera house." *Journal of Personality and Social Psychology* no. 75 (1):198-210.

———. 1998b. "Masterpieces in music and literature: Historiometric inquiries." *Creativity Research Journal* no. 11 (2):103-110.

Stark, Rachel E. 1978. "Features of infant sounds: The emergence of cooing." *Journal of Child Language* no. 5 (3):379-390.

Strunk, Oliver. 1945. "Intonations and Signatures of the Byzantine Modes." *Musical Quarterly* no. 31 (3):339-355.

Tao, Jianhua, Yongguo Kang, and Aijun Li. 2006. "Prosody conversion from neutral speech to emotional speech." *Audio, Speech, and Language Processing, IEEE Transactions on* no. 14 (4):1145-1154.

Taruskin, Richard. 1995. *Text and act: Essays on music and performance*. NY: Oxford University Press.

Titze, Ingo R. 1973. "The human vocal cords: a mathematical model." *Phonetica* no. 28 (3-4):129-170.

———. 1994. *Principles of voice production*. Englewood Cliffs: Prentice Hall.

Trainor, Laurel J., and Beth Adams. 2000. "Infants' and adults' use of duration and intensity cues in the segmentation of tone patterns." *Perception & Psychophysics* no. 62 (2):333-340.

Troelsgård, Christian. 2011. *Byzantine Neumes: A New Introduction to the Middle Byzantine Musical Notation*. København (Copenhagen): Museum Tusculanum Press.

Ververidis, Dimitrios, and Constantine Kotropoulos. 2006. "Emotional speech recognition: Resources, features, and methods." *Speech communication* no. 48 (9):1162-1181.

Welch, Graham F. 1994. "The assessment of singing." *Psychology of Music* no. 22 (1):3-19.

West, Martin Litchfield. 1992. *Ancient greek music*. NY: Oxford University Press.

Wicks, Sammie Ann. 1989. "A Belated Salute to the'Old Way' of 'Snaking' the Voice on its (ca) 345th Birthday." *Popular Music* no. 8 (1):59-96.

Wolfe, Joe. 2002. Speech and music, acoustics and coding, and what music might be 'for'. Paper read at Proceedings of the 7th International Conference on Music Perception and Cognition, at Sydney. Retrieved on June 3, 2007 from http://www.phys.unsw.edu.au/%7Ejw/ICMPC.pdf.

———. 2007. "Speech and Music: Acoustics, Signals and the Relation between them." *Proceedings of ICoMCS December*:176.

Wolfe, Joe, Maëva Garnier, and John Smith. 2009. "Vocal tract resonances in speech, singing, and playing musical instruments." *HFSP journal* no. 3 (1):6-23.

Wolfe, Joe, and Emery Schubert. 2010. "Did non-vocal instrument characteristics influence modern singing?" *Musica Humana* no. 2:121–138.

———. 2012. Stable, quantised pitch in singing and instrumental music: signals, acoustics and possible origins. Paper read at Acoustics 2012 Fremantle: Acoustics, Development and the Environment 21-23 November 2012, at Perth, WA

Yip, Moira. 2002. *Tone*. Cambridge, UK: Cambridge University Press.

Youell, Amber Lynne. 2012. *Opera at the Crossroads of Tradition and Reform in Gluck's Vienna*, Columbia University.

Yurchenco, Henrietta. 1995. "" Blues Fallin'Down Like Hail" Recorded Blues, 1920s-1940s." *American Music* no. 13 (4):448-469.

---

[i] It is also true, of course, that phonemes are well suited to being produced by the human voice, and this statement reflects the order of effect and cause.

[ii] To exemplify the confusion in notating portamento, see Potter's discussion of Norris Croker's 1895 *Handbook for Singers*, in which Croker "follows his basic definition with a warning that the same sign is used for a phrase mark, a bind (what we now call a tie), and a slur (his preferred term for portamento), and that the student has to use 'careful judgement' as to which is which." (Potter 2006, , p. 543).

[iii] Our interpretation has been influenced by comments we received in an earlier draft from two anonymous reviewers, to whom we are most grateful.

[iv] In Porter's musical dictionary of 1834 the link between speech and singing is mentioned under the entry 'Portamento di voce', where it states that "The slide is a grace of much a simplicity and beauty, evidently drawn from nature: it expresses the most tender and affectionate emotion. We hear it in those little gusts of passion which mothers use in caressing their infants; and in the language of nature, it is one of our most endearing tones. This effect is produced by a gradual raising or falling of the voice from any given tone to the next, in one unbroken stream of sound." (Porter 1834, , p. 317)

[v] We were unable to find sources that indicated both how frequently and where *portamento* should be used. References to infrequent use are usually vague. One of the more extreme examples was published during the early electric recording era by Harry Gregory Hast, a singing teacher and publisher of a book on the art of singing (Hast 1925), who in 1929 was invited to answer a series of questions about what makes a singer 'vocal' versus 'unvocal'. His contribution amounted to a general invective on singers, and of portamento he wrote: "What an epidemic of this disease is raging at the moment! Some singers appear to be incapable of singing a descending interval without a horrible slur. It is the cheapest and easiest performed stunt in the repertoire and will reduce temperament to gush. No good singer should indulge in it except as an embellishment—never as a habit. Once, when we were discussing this, Sims Reeves laid down this rule, 'One portamento in a song, never more.' When I told Santley that, he said, in his usual gruff manner, 'And that's once too often.'" (Hast 1929, , p. 252)

[vi] Some writers have ridiculed the use of introspection for gathering data on musical parameters. In a study by Maher of the use of vibrato during *portamento*, such an approach is explicitly rejected because the musician has automated the qualities under investigation and therefore does not necessarily have conscious access to them. Maher exclaims: "Asking a trained singer to explain vibrato behavior during a portamento transition is analogous to asking a runner whether the right foot touches the ground before the left foot is lifted when running at full speed. In both cases the individual does not actively control that physical parameter; it happens automatically as a side effect of singing (or running) and is therefore not introspectable. Thus we must somehow observe and analyze the natural vibrato behavior indirectly" (Maher 2008, p. 21). Obviously there are arguments both ways—the lone musician may not be interested or aware of the generalizability of their perception to a broader group of listeners, but by the same token to argue that they are unable to report what they hear is equally problematic and would be highly limiting for their research colleagues. Teachers are able to make highly perceptive introspections, and Garcia, a renown singing teacher in his time, also had a sense of the parameters involved in *portamento* by distinguishing different kinds of affects from different species of *portamento* as demonstrated in the quote present in §Hypothesis generation. At the very least, such introspections from musicians may still be highly valuable in generating hypotheses.

Biographies

Emery Schubert is an Associate Professor in the School of the Arts and Media at UNSW Australia. His research training consists of a PhD in Music Psychology in the School of Music and Music Education at UNSW Australia, as well as Bachelor degrees in Arts and in Engineering, and a Diploma in Education. His research interests are in interdisciplinary approaches to understanding music, in particular emotional and expressive aspects. He has published in the subdisciplines of music psychology, musicology and education, and is an editor of the forthcoming OUP volume, Expressiveness in music performance: Empirical approaches across styles and cultures.

Joe Wolfe is a Professor in the School of Physics at UNSW Australia. He has a BSc in physics from the University of Queensland, a PhD in applied maths from the Australian National University, and a BA in French from UNSW Australia, acquired while he was teaching there but collaborating with French colleagues. He leads a lab specialising in the acoustics of the voice, the ear and musical instruments. He was won awards from the acoustical societies of America, Australia and France.