Evidence of the development of vocal pitch matching ability in children

Graham F Welch

Institute of Education, University of London, UK

Introduction

Research suggests that singing behaviours are subject to developmental processes in which individual neuropsychobiological potentiality is shaped (nurtured and/or hindered) by learning experiences within socio-cultural contexts (Welch, 2007; in press; Knight, 2009). Although singing is commonplace, it is also marked by cultural diversity, with development related to opportunity (e.g. Mang, 2007), the prosodic features of indigenous languages (Azechi, 2008) and also the dominant characteristics of the local musical soundscapes (Welch et al, 1997; Welch, 2006a; 2006b; in press).

In many parts of the world, the ability to sing is seen as a mark of an individual’s underlying musicality (cf. Sloboda et al, 2005). Consequently, those individuals whose singing development has been hindered in some way are often labelled (including self-labelled) in some absolutist sense under a bi-polar categorisation of ‘can’/’cannot’ sing, with variations in their ascribed musical identity as a ‘non-singer’, ‘tone-deaf’, or ‘tone-dumb’ being found in virtually all cultures. Yet, as mentioned above, contrary evidence from developmental and neurological studies continues to emerge that singing and musical behaviours are context bound and susceptible to improvement with appropriate experience which can be informal as well as formal (e.g. Brown et al, 2004; Koelsch et al 2005; Mang, 2006; 2007; Dalla Bella et al, 2007; Kleber et al, 2007; Fuchs et al, 2007; Mithen & Parsons, 2008; Stewart & Williamson, 2008; Welch et al, 2008; see Welch, 2006a for review).

Furthermore, the recent wealth of studies into the neurosciences and music (cf. Avanzini et al., 2003; 2005) continue to amass evidence of the multi-sited representation of musical behaviours in various regions of the brain, including singing (Kleber et al., 2007). These and related studies also indicate that there are various other-than-musical benefits that can accrue for the individual from engaging in musical (including singing) activity, such as related to physical and psychological health and well-being (Clift & Hancox, 2001; Clift et al., 2007; Kreutz et al., 2004; Welch, 2005), social skill development and social inclusion (Odena, 2007; Portowitz et al., 2008) and cognitive development (Schlaug et al., 2005).

**Early Childhood and Pre-School**

Singing development pre-school is characterised by an increasing interaction with the sounds of the experienced maternal culture. This interaction is reflected in a mosaic of different singing behaviours that are evidenced between the ages of one and five years. They relate to the young child’s acquisitive, playful, creative and spontaneous nature as they engage with and make sense of their “local” musical world. The variety of vocalisation includes: two-year-olds’ repetition of brief phrases with identifiable rhythmic and melodic contour patterns (Dowling, 1999), three-year-olds’ vocal interplay between spontaneous improvisation and selected elements from the dominant song culture, termed “pot-pourri” songs (Moog, 1976), and “outline songs” (Hargreaves, 1996) in which the nature of the figurative shape of the sung melodic contour (its “schematic” contour) is thought to reflect the current level of the young child’s understanding of tonal relationships (Davidson, 1994).

There is evidence of increasing sophistication and complexity in relation to the learning of songs from the dominant culture by young children (and see later for developmental models by Rutkowski, 1997; Welch, 2002). However, the path of development is not necessarily linear for any particular individual. In a USA study of the spontaneous singing of two-year-olds’ first songs, for example, there is evidence that “phrases are the initial musical units” (Davidson, 1994, p. 117). Such phrases are characterised by limited pitch range, a certain disjunction of key/tonality and a descending contour. In contrast, recent Italian data of two- to three-year-old children indicate that some young children appear to be much better at imitating a complete melody modelled by their mother (and also by a specialist course tutor) than in matching individual phrases of the same song (Tafuri, 2009).
For the youngest children, the boundaries between singing and speaking may be blurred, or at least ambiguous to the adult listener, and are related to the dominance of a particular contour schema (Davidson, 1994) as well as to the influence of the mother tongue. For example, a longitudinal study in Canada of young girls aged 18 to 38 months from monolingual and bilingual backgrounds reported that “intermediate vocalisations” (a type of vocal behaviour at the boundary between speech and song) were more prevalent in Mandarin and Cantonese-speaking children than in English-speaking children (Mang, 2000/1).

The First Years of Schooling

It is common for a diverse range of singing abilities to be exhibited by children on entry to compulsory schooling. Within this diversity, it is necessary to distinguish between (i) children’s (developing) skill in the performance of a taught song (cf Rutkowski, 1990, 1997; Welch, 1986, 1998, 2002; Welch et al, 1996, 1997, 1998) and (ii) children’s ability to invent songs (Davies, 1986, 1992, 1994). As with pre-school singing behaviours, context and culture are also factors (cf Rutkowski & Chen-Haftek, 2000; Mang, 2003).

With regard to the first of these categories concerning the skilled performance of a taught song, two major USA and UK studies have drawn on developmental theories to propose phased models of singing development (Rutkowski, 1997; Welch, 1998 – see Figure 1). The USA data (Rutkowski, op.cit.) was generated through systematic evaluation of children’s singing behaviours across a period of over fifteen years. The emergent nine-phase model (which went through several versions) suggests that children progress from speech-like chanting of the song text, to singing within a limited range (“speaking range singer”) to the demonstration of an expanded vocal pitch range that is allied to skilled competency in vocal pitch matching. This model has an affinity with that of another USA-based longitudinal study (Davidson, 1994) that suggests that children’s singing development is linked to a schematic processing of melodic contour. Data from Harvard University’s six-year Project Zero study of children aged between the ages of one and six years indicated five specific levels of pitch development in young children’s singing, expanding from an initial melodic contour scheme with a pitch interval of a third to one that embraced a complete octave.
Within the research literature, children are sometimes reported as being more skilled when copying a sung model if they used a neutral syllable rather than attempting the song with its text (e.g. Levinowitz, 1989). This finding has resonances with data from a three-year longitudinal study of 184 children in their first three years of formal education in ten UK Primary schools (Welch et al., 1996, 1997, 1998). The research provided detailed evidence of how singing behaviours are age, sex and task-sensitive. Over the three years, the participants as a collective appeared to demonstrate little overall improvement when required to match the sung pitches of the criterion songs (two songs were specially taught and assessed each year). However, this singing behaviour was in marked contrast to their ability to learn the words of the songs, which was extremely good, even in their first term of compulsory schooling at age 5. Furthermore, when the pitch elements of the target songs were deconstructed into simpler musical tasks in


1. “Pre-singer” does not sing but chants the song text.
1.5 “Inconsistent Speaking Range Singer” sometimes chants, sometimes sustains tones and exhibits some sensitivity to pitch, but remains in the speaking voice range (usually a3 to c4 [note: the pitch labels have been altered to bring them in line with modern conventions in which middle C = c4, 256 Hz]).
2 “Speaking Range Singer” sustains tones and exhibits some sensitivity to pitch but remains in the speaking voice range (usually a3 to c4).
2.5 “Inconsistent Limited Range singer” waivers between speaking and singing voices and uses a limited range when in singing voice (usually up to f4).
3 “Limited Range Singer” exhibits consistent use of initial singing range (usually d4 to f4).
3.5 “Inconsistent Initial Range Singer” sometimes only exhibits use of limited singing range, but other times exhibits use of initial singing range (usually d4 to a4).
4 “Initial Range Singer’ exhibits consistent use of initial singing range (usually d4 to a4).
4.5 “Inconsistent Singer” sometimes only exhibits use of initial singing range, but other times exhibits use of extended singing range (sings beyond the register lift: b4 and above).
5 “Singer” exhibits use of extended singing range (sings beyond the register lift: b4 and above).


**Phase 1**
The words of the song appear to be the initial centre of interest rather than the melody, singing is often described as ‘chant-like’, employing a restricted pitch range and melodic phrases. In infant vocal pitch exploration, descending patterns predominate.

**Phase 2**
There is a growing awareness that vocal pitch can be a conscious process and that changes in vocal pitch are controllable. Sung melodic outline begins to follow the general (macro) contours of the target melody or key constituent phrases. Tonality is essentially phrase based. Self-invented and ‘schematic’ songs ‘borrow’ elements from the child’s musical culture. Vocal pitch range used in ‘song’ singing expands.

**Phrase 3**
Melodic shape and intervals are mostly accurate, but some changes in tonality may occur, perhaps linked to inappropriate register usage. Overall, however, the number of different reference pitches is much reduced.

**Phase 4**
No significant melodic or pitch errors in relation to relatively simple songs from the singer’s musical culture.

Figure 1: Two independent measures of children’s singing development
which the children were required to match individual pitches, echo melodic contours, or copy small melodic fragments, the children were significantly more pitch accurate, as demonstrated by year-on-year improvements. There were no sex differences in their singing of these three types of deconstructed tasks. Boys and girls were equally successful and demonstrated similar improvements over time. In contrast, when the same boys were faced with the challenge of singing a complete song, their vocal pitch became less accurate and, as a group, they demonstrated little or no improvement in song singing across the three years. Overall, singing competency appeared to be closely related to the nature of the task, with many boys negatively affected in the task of singing a “school” song.

Older Childhood

The latter years of childhood are characterised by a general singing competency for the majority. Relatively few children are reported as singing “out-of-tune” at the age of eleven years (Howard et al, 1994; Welch, 1979; 2002). For example, evidence from a wide range of studies indicates that approximately 30% of pupils aged seven years are reported as being relatively “inaccurate” when vocally matching a melody within a Western cultural tradition. However, this proportion drops to around 4% of the same pupil population by the age of eleven. Within each of these and the intervening age groups, “out-of-tune” boys outnumber girls by a ratio of 2 or 3:1 (Welch, 1979). Culture, however, continues to be significant. Anthropological and ethnomusicological studies, for example, have suggested that young children from the Anang in Nigeria can sing “hundreds of songs, both individually and in choral groups” by the age of 5 (Messinger, 1958: 20), Venda children in South Africa were reported as both learning special children’s songs and composing new songs for themselves (Blacking, 1967), whereas Herati children in Afghanistan tended to focus on the imitation of adult models, with the children (particularly boys) of professional musicians’ families (sazendeh) being immersed in the local music culture and often expected to perform professionally by the age of twelve (Doubleday & Baily, 1995).
Recent UK evidence of the development of vocal pitch matching ability in children

The UK Government has a formal commitment to music, termed its ‘Music Manifesto’ and this includes a National Singing Programme (2007-2011) that has been devised to put ‘group singing at the heart of all primary school musical activity’ (Music Manifesto Report No 2, 2006:8). The programme (called Sing Up, see http://www.singup.org/ for more details) was launched in November 2007 and a team from the Institute of Education, University of London, led by the author, were appointed to undertake a research evaluation of key elements of the programme across its four years.

An initial baseline survey and early impact evaluation were undertaken in 2007-2008 with children drawn from eighty-one schools located across England. Within each school, participant children were drawn primarily from two contrasting age groups, 7-year-olds and 10-year-olds, representing the upper and lower age groups within ‘Key Stage 2’ of the National Curriculum in England. Previous research (e.g. Rutkowski, 1997; Stadler Elmer, 2002; Welch, 1998; 2006a, 2006b; 2007) had demonstrated that clear developmental differences in singing behaviour by age and sex were likely to be evidenced by the selection of these two age groups. Where these target children were in mixed age classes, their classmates were also included.

The initial assessment phase ran from late September 2007 through to February 2008 and was focused on generating some sense of the commonality and diversity of singing behaviours across pupils in English Primary schools. This phase was termed the Year 1 ‘Baseline Assessment’. In this baseline phase, n=3,510 children were assessed from 77 schools. Of these, 10 schools subsequently were visited again, i.e. one visit during the baseline phase and then again between May to July 2008 after a specific Sing Up singing development intervention (called ‘Singing Playgrounds’²) and four schools were visited post-intervention only. Overall, n=394 pupils were assessed in the post-intervention phase. In total, approximately equal numbers of individual singing assessments were made of boys and girls across all phases (1,727 boys and 1,637 girls).

² This particular intervention was termed ‘Singing Playgrounds’ and was provided by members of Ex Cathedra, one of the UK’s leading choir and Early Music ensembles. ‘Singing Playgrounds’ is an educational outreach programme designed to develop children’s musicianship through singing games. Expert adult singers visit school playgrounds and work with older children – called ‘Song Leaders’ – who then lead their peers in singing games in the playground and in class. http://www.singup.org/teachers_and_music_leaders/recipes_for_success/Singing_Playgrounds.php
The research protocol for the assessment of singing and other vocal behaviours drew on established models on singing development from the literature. The protocol included an investigation of each child’s singing behaviour for two well-known song items (either ‘Twinkle, Twinkle’ and ‘Happy Birthday’, or one or other items that the particular child knew well – on advice from the teacher – if these particular songs were unknown). The last of these three elements was assessed against two established rating scales (Rutkowski, 1997; Welch, 1998) (see Figure 1 above). Previous research (Mang, 2006) had demonstrated that the two scales could be used alongside each other to investigate complimentary aspects of singing development. The Rutkowski (1997) scale is a measure of singing voice development, whereas the Welch (1998) scale assesses vocal pitch-matching development.

Children were visited in their schools where their singing was assessed in a quiet space. Each child was taken through the assessment protocol, normally being tested individually within a small group that was drawn from the class. This allowed the other members of the group to observe and see what was required as this had been shown previously to be an appropriate method of accessing better quality responses than individual testing alone (cf Plumridge, 1972). To avoid the effects of vocal modelling, no starting pitch was given for the song items and, although the member of the research team provided verbal encouragement to the child, they did not offer any sung prompt (cf as advised by Mang, 2006). All children completed the singing assessments and none were excluded from the study. Participants’ responses were noted onto individual assessment forms (see Welch et al, 2008 for an example) and data were subsequently entered for collation and analysis into a bespoke data entry form that was connected to a structured query language (SQL) based database. Data analyses included the transformation of the singing scores from the two ratings scales into a combined ‘normalised score’ for which 100% meant that the individual child had scored the highest developmental ratings on both scales.

There were clear, statistically significant, differences evidenced in children’s vocal pitch matching (see Figure 2 and Welch et al 2009[a][b] for details). For example, analyses of the baseline data (n=3,510) indicate that there were age and sex differences in singing development. In general, (a) older children (age 10+) tended to be rated as significantly more developed on both rating scales than their younger peers (age 7+) and (b) girls tended to be rated as more developed than boys in each age group. This is evidenced on
both rating scales (Rutkowski; and Welch) separately as well as in the combined normalised score for each age group (Figure 2).

Figure 2: Average normalised singing development ratings (representing a combination of the Rutkowski and Welch scores) for n=2970 children (n=1472 children aged 7+ years and n=1498 aged 10+ years)

Figure 3: Average normalised singing development ratings (representing a combination of the Rutkowski and Welch scores) post-intervention for n=394 children
A similar pattern is evidenced in the post-intervention data (n=394), but with significantly higher normalised mean scores (see Figure 3). Furthermore, although across the whole data set boys tended to have a lower singing development rating than girls, both at age 7+ and age 10+, boys did make significant progress in their singing development as a result of their participation in the ‘Singing Playgrounds’ activities. There was also evidence that the post-intervention children’s comfortable singing ranges had increased by three semitones in the space of a few months.

When pupil ethnicity (using the official classification from the Ministry of Education [DCSF]) is considered within and across the two data collection phases (baseline and post-intervention), Asian pupils tended to score significantly lower in their singing development ratings compared to their White and Black peers (who tended to be rated similarly). However, notwithstanding these statistical differences between ethnicities, all three major groups had significantly higher normalised singing ratings in their post-intervention assessment data (Figure 4). In the case of the Asian pupils, the post-intervention scores were also much higher than that for the White and Black pupils at baseline. This suggests that all ethnic groups are equally capable of improving their singing abilities and, at a school level, of being at an equivalent developmental level.

![Figure 4: Estimated marginal means for normalised singing scores by ethnicity](image-url)
Conclusion

The latest UK data (now extended with a further 4,500 pupil assessments in 2008-2009 and totalling 8,000 pupils from over 150 Primary schools) continues to confirm these findings. Singing is subject to a development process in which vocal pitch matching (a) improves with age and (b) is subject to accelerated development in an appropriately nurturing environment. Improvement is possible for all groups, including those that tend to have somewhat less advanced singing developmental profiles, such as boys (compared with girls) and Asian children (compared with other ethnic groupings amongst our participants). We also found clear school differences, with schools in very similar socio-economic contexts having quite different singing development profiles. However, there was no one type of school that predominated. More advanced singers could be found in inner city, suburban or rural areas, in both relatively rich and poor neighbourhoods, and with diverse mixes in the school populations regarding gender and ethnicity. What seems to make a difference is the relative importance attached to singing by the school’s senior management team. In schools where singing was seen as important, we tended to find more advanced singing development being evidenced. Singing development should be considered as a normal feature of children’s musical engagement with the world around them, particularly when they are provided with new positive singing experiences.

Acknowledgements

The research team consisted of the author, working closely with a core team of Evangelos Himonides, Jo Saunders and Ioulia Papageorgi, and supported by Tiija Rinta, Costanza Preti, Claire Stewart, Maria Vraka, Jennifer Lani and Joy Hill. The team wish to thank Maurice Walsh, Senior Vocal Tutor with Manchester Music Service; Ula Weber of Ex Cathedra; Dr Penelope Harnett, University of West of England; Dr Liz Mellor, York St John University; and Sarah Kekus and Edward Milner of the Sage Gateshead for their invaluable support in identifying participant schools. We are also extremely grateful to all the schools (pupils, teachers and headteachers) for their time and commitment to participate in this national research activity.

References


singing or listening on secretory immunoglobulin A, cortisol and emotional state. 
Journal of Behavioral Medicine, 27, 623-635.


