Advancing Interdisciplinary Research in Singing through a Shared Digital Repository

Annabel Cohen

University of Prince Edward Island, 550 University Ave, Charlottetown, PE, Canada C1A 4P3
acohen@upei.ca

The domain of research in singing encompasses numerous disciplines, countless styles, and many lifespan stages of skill development. The goal of developing a comprehensive understanding of this domain would benefit from a vast digital repository for storing, accessing, and annotating recordings of singing in all its manifestations. The proposal parallels in part the Child Language Data Exchange System (CHILDES) developed and maintained by Brian MacWhinney (http://childes.psy.cmu.edu/) and used by hundreds of psycholinguistics researchers worldwide. A prototype of a digital repository for singing (http://vre.upei.ca/airs) is under development at the University of Prince Edward Island as part of a major infrastructure project directed toward determining the role of media in education. The database will support and be supported by an international team aiming to Advance Interdisciplinary Research in Singing (AIRS) on the following themes: (1) acquisition of singing (2) comparison of singing and speech acquisition (3) singing pedagogy (4) singing and intercultural understanding (5) singing and intergenerational understanding (6) singing and well-being. The breadth of the work of Johan Sundberg, an honoured member of this team, encompasses these themes and exemplifies the international, interdisciplinary scope of the initiative.

1   Introduction

The acquisition of speech is recognized as one of the most extraordinary aspects of human development. The acquisition of singing is no less remarkable, but unlike speech acquisition, it has received less attention.

Singing and speaking are both fundamental modes of vocal expression entailing complex rule-governed acoustic patterning (Sundberg, 1987). It is often said that children sing before they can speak (Stadler Elmer, 2000; Welch, 1994). They must learn to distinguish singing from speaking, a complex problem of auditory scene analysis (Bregman, 1990). Whereas singing is associated with art, leisure, and worship, speaking is associated with pragmatic communication. Yet, the distinctions between their functions are not clear cut. In its use of words, song entails speech and in its use of intonation, speech entails melody. The overlap between singing and speaking should not mean that what is known about speech acquisition applies to singing. While an understanding of singing may provide insights into the mechanisms of language acquisition, it is critical that we examine singing in its singular context. A comprehensive investigation of singing should address such questions as: how do linguistic, musical and social environments influence the acquisition of singing? What, if any, are the universal aspects of singing acquisition? How do acquisition of singing and speaking differ? Is there a sensitive period during which singing can best be acquired? What are the best methods of teaching singing? And can singing in early life enhance lifelong intercultural understanding, intergenerational understanding, or individual and societal well-being?

Linguistic ability is normally distributed across a population; in contrast the ability to sing might seem to be reserved for those with special talent (Pinker, 1997, p. 529). Yet, data reported by Dalla Bella et al. (2007) revealed that the majority of persons can sing a melody quite accurately. In short, they showed that everyone is equipped to sing. It is true that speaking is essential to survival in society whereas singing, at least in the western world, appears less essential. Yet every culture in the world has some form of singing. Singing may play many significant positive roles in society, roles whose potential may be underexploited in our own society and beyond.

Little is known about the *acquisition* of singing in human development in contrast to the vast amount of information accumulating in other areas of the musical sciences in recent years, the enormous advances in the acoustics of the singing voice for which Sundberg (1987) laid the foundation, and the acquisition of the ability to speak.

The acquisition of singing remains one of the last relatively unexplored frontiers in the growing field of music cognition. The understanding of the acquisition of speech greatly advanced as a result of Brian MacWhinney and Catherine Snow's co-founding of the Child's Language Data Exchange System (CHILDES) database for conversation research in 1984. The database makes use of text and symbols to transcribe speech utterances collected by individual researchers. Codes or rules for data descriptors (metadata), developed by MacWhinney, describe the words or utterances and enable researchers worldwide to flexibly search the data base. CHILDES was not designed for singing because encoding was limited to symbols (letters, linguistic, etc.) rather than lower acoustic level. Research on singing requires the representation of acoustic data. Thanks to recent technological developments, digital libraries for acoustic (sound wave) information of singing have become feasible. The iPod and other portable audio playback devices are examples of relatively small digital music libraries. (CHILDES and related TALKBANK more recently accept acoustical data, but the entire enterprise is directed to language transcription.)

We have begun to develop a comparable digital database for singing. It is part of a large digital library known as CMTC, an infrastructure project aiming to determine the best use of media in the context of Culture, Multimedia, Technology and Cognition. A collection within this digital library is focused on singing. This collection is referred to as AIRS (Advancing Interdisciplinary Research in Singing). Within AIRS resides the collection known as CHIMES (Children's International Music Exchange System). The CHIMES database aims to be analogous in many ways to the CHILDES database for language, in that entries include longitudinal and cross-sectional data on singing acquisition, and associated transcriptions, analyses, and representations.
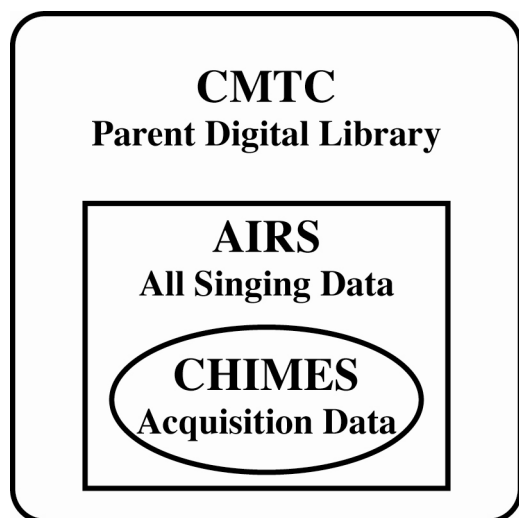
Fig. 1: Relation of the CHIMES, AIRS, and CMTC digital library repositories (see text) for explanation.

There is more to CHIMES, the proposed digital library for singing, however than simply storing acoustical records. As with CHILDES, standards must be developed for materials allowed into the research database. A system for describing the entries (metadata) must be developed so that recordings can be organized for storage and easy retrieval, e.g., with respect to the age, gender, and musical, linguistic and educational background of the singer, the communication setting, the cultural and historical context, and the quality of the recording. Protocols must also be developed for transcribing, analyzing, and annotating raw data. The success of the system lies also in depositing results of analysis of the data and linking them to the raw data so that other researchers may benefit from accumulating results and build on them. Singing is an activity potentially involving much of the entire human body (Sundberg, 1987). From the age of the newborn onwards, it is part of social interaction. Video records of the singing activity together with audio records provide critical information describing singing within a social context, and so not only audio but also some video data would ideally be collected.

Exploiting the Digital Library of CMTC. The proposed CHIMES digital archive will exploit a digital library that is the chief infrastructure for an existing Canada Foundation for Innovation (CFI) project entitled the "Institute for Interdisciplinary Research in Culture, Multimedia, Technology and Cognition (CMTC; see http://vre.upei.ca/cmtc). CMTC research aims to determine how multimedia can best serve education in a cultural context. The capacity of the CMTC digital library will enable the study of the role of culture in multimedia education, by storing cultural artefacts, of which singing in its many forms and contexts provides examples. The digital library will also store audiovisual teaching materials or "learning objects" to enable a comparison of teaching methods with

and without various kinds of multimedia, and it will store results of studies so as to share findings among the researchers. The study of how best to teach singing, the study of the mind that enables singing, and the development of multimedia examples for the teaching of either singing or culture linked to a song are all consistent with the CMTC goals, and justify the use of the CMTC digital library to support CHIMES research.

Developmental Cognitive Science of Music. Extensive research over the last three decades has been directed to defining mental representations of tonal (western) musical structure. Studies have shown that rules of music theory describe the relations among tones represented by listeners to music (Cohen, 1991, 2005; Cuddy, Cohen & Mewhort, 1981; Halpern, 1984; Krumhansl,1990). Tone sequences that obey rules of musical harmony are more memorable than those that do not. Infants prefer regular musical patterns (Trehub, 2001), major triads (Cohen, Thorpe & Trehub, 1987), and tonality (Trehub, Cohen, Thorpe & Morrongiello, 1986). From infancy onward, music lacking tonal structure (i.e., atonal, serial, or 12-tone), challenges organization for the majority of western listeners. Yet the western tonal system is not culturally universal. General cognitive principles such as similarity detection must also be considered (Deliège, 2001). Cross-cultural investigations of singing can be used to test theories about the role of innate versus experiential influences on the development of musical cognitive processes. The role of language in acquisition of singing and in music cognitive development is an important consideration. More accurate pitch of singing has been shown among speakers of a tone language (Mang, 2006). Cohen (2000), and Trainor (2004) also discuss the possibility for a critical period for acquisition of music, a notion borrowed from psycholinguistics (cf., Thomas & Johnson, 2008).

Developmental Psycholinguistics. Psycholinguistics focuses on the acquisition of language. Much data have been collected on the emergence of particular words, word patterns, grammatical errors and correct constructions across development in different cultures. Data from 26 languages reside in the CHILDES database and have served hundreds of researchers leading to over 1500 publications (Macwhinney, 2000; Berko & Thompson, 2002). These data can be used to test theories of language acquisition. One theory inspired by the eminent linguist Noam Chomsky argues that basic linguistic structure is present at birth, but exposure to a language is needed to set linguistic parameters appropriate to the particular language environment. A key notion, originating with Penfield promoted by Lenneberg (1967) and endorsed by Chomsky suggests a critical or sensitive period before adolescence beyond which acquisition of first or second language is difficult. A contrasting approach emphasizes experience and exposure to the statistical properties of language as determining what rules of language are acquired (Saffran, 2004). Yet a more social approach focuses on human interaction and feedback in language development, including motivational factors (Gros-Louis et al., 2006). These perspectives can in theory also be applied to the acquisition of singing.

Vocal Pedagogy. Knowledge of the normal pattern of acquisition of singing can assist in improving the teaching of singing and in measuring student progress. It

also may help to resolve past controversies about pedagogy, such as those concerning the best method for teaching sight-singing (Demorest, 2001). This knowledge of normal song acquisition may also be used to help develop guidelines for appropriate use of multimedia for teaching singing, including audiovisual tools. The multimedia educational expert Richard Mayer (2005) defines multimedia in terms of text/narration and visual images, and he argues that people learn better by text and images than by text alone. But how does this translate to enhancement of teaching of a skill like singing? The AIRS multimedia database of songs and audiovisual recordings of singers in cultural context can be used to enhance teaching, and to develop techniques for teaching singing, particularly singing of songs of other cultures. It eventually can offer both a resource of songs, and images of their cultural context.

Cross-cultural understanding. The consequences of learning songs of another culture early in life may have far reaching social implications. Research in Portugal showed that children in a majority culture who learned songs of a minority culture exhibited improved attitudes toward the minority culture (Sousa, Neto, & Mullet, 2005). Prejudice and negative attitudes are natural responses to difference. In multicultural societies like Canada, methods that can assuage prejudice and foster intercultural understanding are valuable for the maintenance of well-being and peace. The AIRS database will be designed so as to provide a repository for singing from people of many cultures and nationalities. Stored videos of singing of songs in their natural context could assist educators in teaching these songs to members of different cultures, so as to potentially foster intercultural understanding. Social psychological models of acculturation will also inform this research (e.g., Berry, 2003).

Intergenerational Educational Research. Frequently the oldest generations in our western societies live in isolation and loneliness, yet they are the human repositories of the cultural traditions of the particular society. Rachel Heydon (in press) of the University of Western Ontario initiated research on intergenerational art education in which children meet regularly with elderly persons living in a home for seniors. She has published guidelines for such activities that can be adapted to intergenerational singing. In some cultures, such as the aboriginal community, elders are the ones who pass down to the children the singing traditions. Thus, an important aspect of singing is the generational context: singing can bridge mixed generations and bring value to lives, preserve cultural traditions through intergenerational singing activities, and reduce children's possible fear of, or disrespect for, senior members of society.

Wellbeing and Singing. Methods for improving and maintaining wellbeing and health are invaluable to society. Singing in a choir has been shown to increase positive affect and enhance immune competence (Kreutz, et al., 2004). Surveys of choir members in various counties have also revealed the value of singing to life satisfaction (Bailey & Davidson, 2002, 2005). Singing practice can lead to mastery, self-expression, and a feeling of competence, leading to competence in other areas of life. Singing may lead to the general educational and health benefits of engagement in music and the arts (Burton, Horowitz & Abeles, 2000).

2. Developing CHIMES and AIRS.

The following steps are necessary to the development of the repository for singing. (1) Collect and review all methods for analyzing singing e.g., Sundberg & Thalén's (2001) method for analysis of musical styles; Lomax's (1968) anthropological/ ethno-musicological Cantometrics; Stadler-Elmer's (2000) musical analysis system for children's song, and Praat's general intonation system (Boersma, 2007). (2) development of greater expertise with the CHILDES database for language research, the metadata that describes data entries, the protocols for data access, analysis and annotation. (3) develop CHILDES-like functions for singing data (metadata, protocols for data access, analysis and annotation). Whereas much of the CHILDES data is at a symbolic word level, the singing data could be analyzed at the acoustic level and at higher levels such as musical notation (particularly for music in western cultures), rhythm, emotional interpretation, musical stylistic features, vocal quality, and cultural context. (4) Develop algorithms for searching the database adapting tools developed for digital library.

The AIRS database. In theory any examples of singing and associated files of information are potential residents of the AIRS repository. Data on singing acquisition or singing ability have been collected by many researchers. Such data are painstaking to obtain, and an individual researcher may have time to analyse only one aspect of it. Yet other researchers could re-use this data, and also compare across singing data sets.

My research has led to two such collections (1) musical scale production data associated with determining the key of 12 Bach Preludes cued by excerpts of the first 4 note-events, the first 4 bars, the first 8 bars or the last 8 bars (Cohen, 1991); (2) data of sight-singing 7-note passages differing in musical structure. Examples of these data sets are being imported to CHIMES for future analyses. The data were used for only one purpose but could now address many other issues. New data (see below) will also be deliberately collected and encoded. International colleagues including Lily Chen Hafteck (USA), Mayumi Adachi (Japan), Stephanie Stadler Elmer (Switzerland), and Kate Stevens (Australia) among others have previously indicated interest in sharing vocal data. In theory the data can be entered from any computer connected to the Internet. Protocols are currently being developed to optimize this process. The data can be viewed and auditioned according to permissions granted. Some examples of singing may be of a personal nature, whereby those involved may be willing to present the audio and video data to researchers but not to the public at large. The restriction on audience is less for audio-only materials.

(2) **Singing Acquisition - Behavioral research.** Sample data on the progression of singing with age in particular musical and language contexts is needed. Much can be learned initially from careful data collection from a few children (e.g., Brown, 1972). Currently 4 children in Prince Edward Island are being "interviewed" for singing at each of 3 age groups (3, 5, and 7 years) as well as a control group of 4 adults (2 trained vocalists and 2 without special training). For each participant, "singing interviews" are being videorecorded at approximately monthly intervals 6 months (initially, and ideally longer). Each interview will

allows for (a) free-form improvisation as well as (b) fixed repetition of two model or criterion songs (*Frère Jacques,* known in many languages, and *We are one* an international anthem based on the *Earth Charter*) and (c) repetition of specific musical elements (e.g., short melodic patterns). The acoustical data is being stored in the CHIMES collection for subsequent analysis by ear (contour, grouping, notation etc), by Praat software (available on the Internet and used in my laboratory) and by the software developed by Stadler Elmer, and will be described according to the CHIMES metadata protocol (including cantometric features). Comparisons will be made within individuals across the three sessions, within age and across age groups. Opportunities to collect similar data in a contrasting culture or cultures are envisioned both within Canada (e.g., nearby Acadian French communities) or cultures worldwide.

The data can be used to test (1) the nativist hypothesis e.g., that musical relations characterizing tonal music are the first to be accurately sung (2) the empiricist hypothesis that the child's musical and linguistic world determines emergent patterns of accurate singing. Inclusion of cross-cultural comparisons will also assist in determining commonalities and differences between singing and language. The data will address the questions: When does singing become differentiated from speech (although infants younger than our current children will be needed for this)? What is the role of exposure to the music and language of one's culture on the acquisition of singing or speech characteristics? Do good models of singing facilitate acquisition of singing or are impoverished models sufficient as has been argued by Chomsky for language? Is the timing of the presentation of models critical (critical period hypothesis)? What is the role of human interaction and play?

The aim will be to develop and test models of singing acquisition covering background in physics and physiology of song production, characterizing the patterned acoustic environment, perception of singing, and the role of human interaction in singing, including motivational and social factors. The goal is to develop comprehensive neuroscientifically plausible models that are synchronic (dealing with many factors at a particular time) and diachronic (dealing with change across the lifespan including issues associated with brain plasticity and sensitive periods during which particular aspects of singing can be best acquired). Diachronicity also includes change in song style across generations. Overall the AIRS database will be able to support research on the 6 themes: acquisition of singing, comparison of singing and language, singing pedagogy, singing and cross-cultural understanding, singing and intergenerational activities, and singing and well-being.

With the aim of providing an integrative approach, each of the six research theme can be explored along two axes. The first axis arises from a synthesis of Welch's (2005) review of singing as communication and Gardner's (1999) theory of multiple intelligences. At first singing would seem to fall squarely in the domain of musical intelligence. However, sensory motor, linguistic, interpersonal, and intrapersonal factors are all important (Welch, 2005). We also add the factor of moral intelligence, which applies to issues of how singing, as a powerful force, is used, or could or should be used to benefit society. The second axis is the biopsychosocial dimension that differentiates individual, cultural-group, and universal influences on human development (Leoung, 2007; Kluckhohn & Murray, 1950). Together, the 3 axes (6 themes, 3 biopsychosocial factors, and 6 intelligences) create a three-dimensional framework for situating the AIRS research. Guided by this framework, the AIRS digital library will help to revolutionize research in singing with the outcome of filling the now relatively empty cells of this framework with new knowledge. While the study of singing may appear as an esoteric exercise, it has much to do with increasing knowledge about intelligence and biopsychosocial factors governing behaviour, and provides a fresh perspective on human development.

3. Outcomes

(1) AIRS: building on the foundation of brilliant research conducted on the analysis of the singing voice (e.g., Sundberg and colleagues) a digital repository for singing will help to revolutionize the way in which multidisciplinary research in singing is carried out. By enabling researchers to share data on singing, advances will be made in our understanding of human mind and development.

(2) new integrative diachronic and synchronic theories of singing, will be developed to account for singing acquisition, comparison with language acquisition, singing pedagogy, cross-cultural considerations, intergenerational considerations, and singing and well-being.

(3) Social relevance: In this age of diminishing natural resources it is promising to think that within each of us is the natural and often untapped resource of our singing voice. Knowing how and when to develop singing to its greater potential within society can enrich lives, increase intercultural and intergenerational understanding, reduce stress, improve health and facilitate education.

Acknowledgments

References

Aitkinson, J. (1998). *The articulate mammal.* N.Y.: Routledge.

Bailey, B. A. & Davidson, J. W. (2005). Effects of group singing and performance for marginalised and middle class singers. *Psychology of Music (POM), 33,* 271-305.

Bailey, B. A. & Davidson, J. W. (2002). Adaptive characteristics of group singing: Perceptions from members of a choir for homeless men. *Musicae Scientiae, 6,* 221-256.

Berry, J.W. (2003). Conceptual approaches to acculturation. In K. Chung, P. Balls-Organista & G. Marin (Eds*.). Acculturation* (pp. 17-37). Washington: American Psych. Association.

Berko, J.B.G. & Thompson, R.B. (2002). Out of the baby book and into the computer: Child language research comes of age. [Rev of MacWhinney (2000) *The CHILDES Project: tools for Analyzing Talk, Vol.1& 2*., Mahweh, NJ: Erlbaum] *APA Rev Books, 47,* 391-394.

Boersma, P. & Weenink, D. (2005) Praat: doing phonetics by computer (Version 4.3.14) [Computer program].//www.praat.org/

Bregman, A. (1990). *Auditory scene analysis*. Cambridge, MA.: MIT.

Brown, R. (1973). *A first language*. London: Allen & Unwin.

Burton, J.,Horowitz, R., & Abeles, H. (2000). Learning in and through the arts. In *Champions of Change: The Impact of the Arts on Learning*. Washington: President's Commission on the Arts and Humanities.

Dalla Bella, S., Giguère, J. F., & Peretz, I. (2007). Singing proficiency in the general population. *Journal of the Acoustical Society of America, 121,* 1182-1189.

Deliège, I. (2001).(Ed). Similarity Perception↔ Categorization↔Cue Abstraction. *Mus Perc.,18.*

Demorest, S. (2001). *Building choral excellence: Teaching sight-singing in the choral rehearsal*. Oxford, UK: OUP.

Gros-Louis, J., West, M., Goldstein, M. H., & King, A. P. (2006). Mothers provide differential feedback to infants' prelinguistic sounds. *Internat J. Behav. Development, 30,* 509-516.

Halpern, A. (1984). Perception of structure in novel music. *Memory & Cognition, 12*, 163-170.

Heydon, R. & B. Daly (forthcoming) *The handbook of intergenerational art activities*.

Kreutz, G., et al. (2004). Effects of choir singing or listening on secretory IgA, cortisol, and emotional state. *Journal of Behavioral Medicine, 27*, 623-636.

Krumhansl, C. (1990). *The cognitive structure of musical pitch*. Oxford, UK: Oxford U. Press.

Lenneberg, E. (1967). *The biological foundations of language.* N.Y.: Wiley.

Lomax, A. (1968). *Folk song style and culture.* New Brunswick, NJ: Transaction

MacWhinney, B. (2000). *The CHILDES Project:Tools for Analyzing Talk*. Mahwah,NJ: Erlbaum

Mang, E. (2006). The effects of age, gender and language on children's singing competency. *British Journal of Music Education, 23*, 161-174.

Mayer, R. (2005). (Ed.) *Cambridge Handbook of Multimedia Learning*. Cambridge, UK: C.U.P.

Penfield, W.& Roberts,W.(1959). *Speech & brain mechanisms*. Princeton, NJ: Princeton U.P.

Pinker, S. (1997). *How the mind works*. NY: Norton.

Sousa, M., Neto, F., & Mullet, E. (2005). Can music change ethnic attitudes among children? *Psychology of Music (POM), 33*, 304-316.

Stadler Elmer, S. (2000).A new method for analyzing and representing singing. *POM,28*,23-42.

Sundberg, J. (1987). The science of the singing voice. DeKalb, IL: Northern Illinois University Press.

Sundberg J & Thalén M (2001) "Describing different styles of singing. A comparison of a female singer's voice source in 'Classical','Pop','Jazz' and 'Blues'", Logopedics, Phoniatrics Vocology 26, 82-93.

Thomas, M. S. C., & Johnson, M. H., (2008). New Advances in understanding sensitive periods in brain development. Current directions in psychological science, 17, 1-5.

Trainor, L. J.(2005). Are there critical periods for musical development? *DevPsyB, 46,*262-278.

Trehub, S. E. (2001). Musical predispositions in infancy. In *The Biological foundations of music*, (eds R.J. Zatorre & I. Peretz), 930, pp. 1-16. NY.: Annals NY Academy Sciences.

Welch, G. (2005). Singing as communication. In D. Miell, R. MacDonald & D. J. Hargreaves (Eds). *Musical Communication* (pp. 239-259). Oxford, UK: Oxford U Press.

Note

This paper was presented as part of an ASA-EAA symposium on Interdisciplinary Research in the Science of Singing: A tribute to the work of Johan Sundberg. The symposium was co-organized by Sten Ternström and Annabel J. Cohen.