Communication of Emotions in Vocal Expression and Music Performance: Different Channels, Same Code?

Patrik N. Juslin and Petri Laukka
Uppsala University

Many authors have speculated about a close relationship between vocal expression of emotions and musical expression of emotions, but evidence bearing on this relationship has unfortunately been lacking. This review of 104 studies of vocal expression and 41 studies of music performance reveals similarities between the 2 channels concerning (a) the accuracy with which discrete emotions were communicated to listeners and (b) the emotion-specific patterns of acoustic cues used to communicate each emotion. The patterns are generally consistent with K. R. Scherer's (1986) theoretical predictions. The results can explain why music is perceived as expressive of emotion, and they are consistent with an evolutionary perspective on vocal expression of emotions. Discussion focuses on theoretical accounts and directions for future research.

Music: Breathing of statues.

Perhaps: Stillness of pictures. You speech, where speeches end.

You time, vertically poised on the courses of vanishing hearts.

Feelings for what? Oh, you transformation of feelings into

. . . audible landscape!

You stranger: Music.

—Rainer Maria Rilke, "To Music"

Communication of emotions is crucial to social relationships and survival (Ekman, 1992). Many researchers argue that communication of emotions serves as the foundation of the social order in animals and humans (see Buck, 1984, pp. 31–36). However, such communication is also a significant feature of performing arts such as theater and music (G. D. Wilson, 1994, chap. 5). A convincing emotional expression is often desired, or even expected, from actors and musicians. The importance of such artistic expression should not be underestimated because there is now increasing evidence that how people express their emotions has implications for their physical health (e.g., Booth & Pennebaker, 2000; Buck, 1984, p. 229; Drummond & Quah, 2001; Giese-Davis & Spiegel, 2003; Siegman, Anderson, & Berger, 1990).

Two modalities that are often regarded as effective means of emotional communication are vocal expression (i.e., the nonverbal aspects of speech; Scherer, 1986) and music (Gabrielsson & Juslin, 2003). Both are nonverbal channels that rely on acoustic signals for their transmission of messages. Therefore, it is not surprising

that proposals about a close relationship between vocal expression and music have a long history (Helmholtz, 1863/1954, p. 371; Rousseau, 1761/1986; Spencer, 1857). In a classic article, "The Origin and Function of Music," Spencer (1857) argued that vocal music, and hence instrumental music, is intimately related to vocal expression of emotions. He ventured to explain the characteristics of both on physiological grounds, saying they are premised on "the general law that feeling is a stimulus to muscular action" (p. 400). In other words, he hypothesized that emotions influence physiological processes, which in turn influence the acoustic characteristics of both speech and singing. This notion, which we refer to as *Spencer's law*, formed the basis of most subsequent attempts to explain reported similarities between vocal expression and music (e.g., Fónagy & Magdics, 1963; Scherer, 1995; Sundberg, 1982).

Why should anyone care about such cross-modal similarities, if they really exist? First, the existence of acoustic similarities between vocal expression of emotions and music could help to explain why listeners perceive music as expressive of emotion (Kivy, 1980, p. 59). In this sense, an attempt to establish a link between the two domains could be made for the sake of theoretical economy, because principles from one domain (vocal expression) might help to explain another (music). Second, cross-modal similarities would support the common—although controversial—hypothesis that speech and music evolved from a common origin (Brown, 2000; Levman, 2000; Scherer, 1995; Storr, 1992, chap. 1; Zucker, 1946).

A number of researchers have considered possible parallels between vocal expression and music (e.g., Fónagy & Magdics, 1963; Scherer, 1995; Sundberg, 1982), but it is fair to say that previous work has been primarily speculative in nature. In fact, only recently have enough data from music studies accumulated to make possible a systematic comparison of the two domains. The purpose of this article is to review studies from both domains to determine whether the two modalities really communicate emotions in similar ways. The remainder of this article is organized as follows: First, we outline a theoretical perspective and a set of predictions. Second, we review parallels between vocal expression and music performance regarding (a) the accuracy with which

different emotions are communicated to listeners and (b) the acoustic means used to communicate each emotion. Finally, we consider theoretical accounts and propose directions for future research.

An Evolutionary Perspective

A review needs a perspective. In this overview, the perspective is provided by evolutionary psychology (Buss, 1995). We argue that this approach offers the best account of the findings that we review, in particular if the theorizing is constrained by findings from neuropsychological and comparative studies (Panksepp & Panksepp, 2000). In this section, we outline theory that serves to support the following seven guiding premises: (a) emotions may be regarded as adaptive reactions to certain prototypical, goal-relevant, and recurrent life problems that are common to many living organisms; (b) an important part of what makes emotions adaptive is that they are communicated nonverbally from one organism to another, thereby transmitting important information; (c) vocal expression is the most phylogenetically continuous of all forms of nonverbal communication; (d) vocal expressions of discrete emotions usually occur in similar types of life situations in different organisms; (e) the specific form that the vocal expressions of emotion take indirectly reflect these situations or, more specifically, the distinct physiological patterns that support the emotional behavior called forth by these situations; (f) physiological reactions influence an organism's voice production in differentiated ways; and (g) by imitating the acoustic characteristics of these patterns of vocal expression, music performers are able to communicate discrete emotions to listeners.

Evolution and Emotion

The point of departure for an evolutionary perspective on emotional communication is that all human behavior depends on neurophysiological mechanisms. The only known causal process that is capable of yielding such mechanisms is evolution by natural selection. This is a feedback process that chooses among different mechanisms on the basis of how well they function; that is, function determines structure (Cosmides & Tooby, 2000, p. 95). Given that the mind acquired its organization through the evolutionary process, it may be useful to understand human functioning in terms of its adaptive significance (Cosmides & Tooby, 1994). This is particularly true for such types of behavior that can be observed in other species as well (Bekoff, 2000; Panksepp, 1998).

Several researchers have taken an evolutionary approach to emotions. Before considering this literature, a preliminary definition of emotions is needed. Although emotions are difficult to define and measure (Plutchik, 1980), most researchers would probably agree that emotions are relatively brief and intense reactions to goal-relevant changes in the environment that consist of many subcomponents: cognitive appraisal, subjective feeling, physiological arousal, expression, action tendency, and regulation (Scherer, 2000, p. 138). Thus, for example, an event may be appraised as harmful, evoking feelings of fear and physiological reactions in the body; individuals may express this fear verbally and nonverbally and may act in certain ways (e.g., running away) rather than others. However, researchers disagree as to whether emotions are best conceptualized as categories (Ekman, 1992),

dimensions (Russell, 1980), prototypes (Shaver, Schwartz, Kirson, & O'Connor, 1987), or component processes (Scherer, 2001). In this review, we focus mainly on the expression component of emotion and adopt a categorical approach.

According to the evolutionary approach, the key to understanding emotions is to study what functions emotions serve (Izard, 1993; Keltner & Gross, 1999). Thus, to understand emotions one must consider how they reflect the environment in which they developed and to which they were adapted. On the basis of various kinds of evidence, Oatley and Jenkins (1996, chap. 3) suggested that humans' environment of evolutionary adaptedness about 200,000 years ago was that of seminomadic hunter–gatherer groups of 10 to 30 people living face-to-face with each other in extended families. Most emotions, they suggested, are presumably adapted to living this kind of way, which involved cooperating in activities such as hunting and rearing children. Several of the activities are associated with basic survival problems that most organisms have in common—avoiding predators, finding food, competing for resources, and caring for offspring. These problems, in turn, required specific types of adaptive reactions. A number of authors have suggested that such adaptive reactions were the prototypes of emotions as seen in humans (Plutchik, 1994, chap. 9; Scott, 1980).

This view of emotions is closely related to the concept of *basic emotions*, that is, the notion that there is a small number of discrete, innate, and universal emotion categories from which all other emotions may be derived (e.g., Ekman, 1992; Izard, 1992; Johnson-Laird & Oatley, 1992). Each basic emotion can be defined, functionally, in terms of an appraisal of goal-relevant events that have recurred during evolution (see Power & Dalgleish, 1997, pp. 86–99). Examples of such appraisals are given by Oatley (1992, p. 55): happiness (subgoals being achieved), anger (active plan frustrated), sadness (failure of major plan or loss of active goal), fear (self-preservation goal threatened or goal conflict), and disgust (gustatory goal violated). Basic emotions can be seen as fast and frugal algorithms (Gigerenzer & Goldstein, 1996) that deal with fundamental life issues under conditions of limited time, knowledge, or computational capacities. Having a small number of categories is an advantage in this context because it avoids the excessive information processing that comes with too many degrees of freedom (Johnson-Laird & Oatley, 1992).

The notion of basic emotions has been the subject of controversy (cf. Ekman, 1992; Izard, 1992; Ortony & Turner, 1990; Panksepp, 1992). We propose that evidence of basic emotions may come from a range of sources that include findings of (a) distinct brain substrates associated with discrete emotions (Damasio et al., 2000; Panksepp, 1985, 2000, Table 9.1; Phan, Wager, Taylor, & Liberzon, 2002), (b) distinct patterns of physiological changes (Bloch, Orthous, & Santibañez, 1987; Ekman, Levenson, & Friesen, 1983; Fridlund, Schwartz, & Fowler, 1984; Levenson, 1992; Schwartz, Weinberger, & Singer, 1981), (c) primacy of development of proposed basic emotions (Harris, 1989), (d) cross-cultural accuracy in facial and vocal expression of emotion (Elfenbein & Ambady, 2002), (e) clusters that correspond to basic emotions in similarity ratings of affect terms (Shaver et al., 1987), (f) reduced reaction times in lexical decision tasks when priming words are taken from the same basic emotion category (Conway & Bekerian, 1987), and (g) phylogenetic continuity of basic emotions (Panksepp, 1998, chap. 1–3; Plutchik, 1980; Scott, 1980). It is fair

to acknowledge that some of these sources of evidence are not strong. In the case of autonomic specificity especially, the jury is still out (for a positive view, see Levenson, 1992; for a negative view, see Cacioppo, Berntson, Larsen, Poehlmann, & Ito, 2000).[1] Arguably, the strongest evidence of basic emotions comes from studies of communication of emotions (Ekman, 1973, 1992).

Vocal Communication of Emotion

Evolutionary considerations may be especially relevant in the study of communication of emotions, because many researchers think that such communication serves important functions. First, expression of emotions allows individuals to communicate important information to others, which may affect their behaviors. Second, recognition of emotions allows individuals to make quick inferences about the probable intentions and behavior of others (Buck, 1984, chap. 2; Plutchik, 1994; chap. 10). The evolutionary approach implies a hierarchy in the ease with which various emotions are communicated nonverbally. Specifically, perceivers should be attuned to that information that is most relevant for adaptive action (e.g., Gibson, 1979). It has been suggested that both expression and recognition of emotions proceed in terms of a small number of basic emotion categories that represent the optimal compromise between two opposing goals of the perceiver: (a) the desire to have the most informative categorization possible and (b) the desire to have these categories be as discriminable as possible (Juslin, 1998; cf. Ross & Spalding, 1994). To be useful as guides to action, emotions are recognized in terms of only a few categories related to life problems such as danger (fear), competition (anger), loss (sadness), cooperation (happiness) and caregiving (love).[2] By perceiving expressed emotions in terms of such basic emotion categories, individuals are able to make useful inferences in response to urgent events. It is arguable that the same selective pressures that shaped the development of the basic emotions should also favor the development of skills for expressing and recognizing the same emotions. In line with this reasoning, many researchers have suggested the existence of innate affect programs, which organize emotional expressions in terms of basic emotions (Buck, 1984; Clynes, 1977; Ekman, 1992; Izard, 1992; Lazarus, 1991; Tomkins, 1962). Support for this notion comes from evidence of categorical perception of basic emotions in facial and vocal expression (de Gelder, Teunisse, & Benson, 1997; de Gelder & Vroomen, 1996; Etcoff & Magee, 1992; Laukka, in press), more or less intact vocal and facial expressions of emotion in children born deaf and blind (Eibl-Eibesfeldt, 1973), and cross-cultural accuracy in facial and vocal expression of emotion (Elfenbein & Ambady, 2002).

Phylogenetic continuity.   Vocal expression may be the most phylogenetically continuous of all nonverbal channels. In his classic book, *The Expression of the Emotions in Man and Animals*, Darwin (1872/1998) reviewed different modalities of expression, including the voice: "With many kinds of animals, man included, the vocal organs are efficient in the highest degree as a means of expression" (p. 88).[3] Following Darwin's theory, a number of researchers of vocal expression have adopted an evolutionary perspective (H. Papoušek, Jürgens, & Papoušek, 1992). A primary assumption is that there is phylogenetic continuity of vocal expression. Ploog (1992) described the morphological transformation of the larynx—from a pure respiratory organ (in lungfish) to

a respiratory organ with a limited vocal capability (in amphibians, reptiles, and lower mammals) and, finally, to the sophisticated instrument that humans use to sing or speak in an emotionally expressive manner.

Vocal expression seems especially important in social mammals. Social grouping evolved as a means of cooperative defense, although this implies that some kind of communication had to develop to allow sharing of tasks, space, and food (Plutchik, 1980). Thus, vocal expression provided a means of social coordination and conflict resolution. MacLean (1993) has argued that the limbic system of the brain, an essential region for emotions, underwent an enlargement with mammals and that this development was related to increased sociality, as evident in play behavior, infant attachment, and vocal signaling. The degree of differentiation in the sound-producing apparatus is reflected in the organism's vocal behavior. For example, the primitive condition of the sound-producing apparatus in amphibians (e.g., frogs) permits only a few innate calls, such as mating calls, whereas the highly evolved larynx of nonhuman primates makes possible a rich repertoire of vocal expressions (Ploog, 1992).

The evolution of the phonatory apparatus toward its form in humans is paralleled not only by an increase in vocal repertoire but also by an increase in voluntary control over vocalization. It is possible to delineate three levels of development of vocal expression in terms of anatomic and phylogenetic development (e.g., Jürgens, 1992, 2002). The lowest level is represented by a completely genetically determined vocal reaction (e.g., pain shrieking). In this case, neither the motor pattern producing the vocal expression nor the eliciting stimulus has to be learned. This is referred to as an *innate releasing mechanism*. The brain structures responsible for the control of such mechanisms seem to be limited mainly to the brain stem (e.g., the periaqueductal gray).

The following level of vocal expression involves voluntary control concerning the initiation and inhibition of the innate expressions. For example, rhesus monkeys may be trained in a vocal operant conditioning task to increase their vocalization rate if each

---

[1] It seems to us that one argument is often overlooked in discussions regarding physiological specificity, namely that this may be a case in which positive results count as stronger evidence than do negative results. It is generally agreed that there are several methodological problems involved in measuring physiological indices (e.g., individual differences, time-dependent nature of measures, difficulties in providing effective stimuli). Given the error variance, or "noise," that this produces, it is arguably more problematic for the no-specificity hypothesis that a number of studies have obtained similar and reliable emotion-specific patterns than it is for the specificity hypothesis that a number of studies have failed to yield such patterns. Failure to obtain patterns may be due to error variance, but how can the presence of similar patterns in several studies be explained?

[2] Love is not included in most lists of basic emotions (e.g., Plutchik, 1994, p. 58), although some authors regard it as a basic emotion (e.g., Clynes, 1977; MacLean, 1993; Panksepp, 2000; Scott, 1980; Shaver et al., 1987), as have philosophers such as Descartes, Spinoza, and Hobbes (Plutchik, 1994, p. 54).

[3] In accordance with Spencer's law, Darwin (1872/1998) noted that vocalizations largely reflect physiological changes: "Involuntary . . . contractions of the muscles of the chest and the glottis . . . may first have given rise to the emission of vocal sounds. But the voice is now largely used for various purposes"; one purpose mentioned was "intercommunication" (p. 89).

vocalization is rewarded with food (Ploog, 1992). Brain-lesioning studies of rhesus monkeys have revealed that this voluntary control depends on structures in the anterior cingulate cortex, and the same brain region has been implicated in humans (Jürgens & van Cramon, 1982). Neuroanatomical research has shown that the anterior cingulate cortex is directly connected to the periaqueductal region and thus in a position to exercise control over the more primitive vocalization center (Jürgens, 1992).

The highest level of vocal expression involves voluntary control over the precise acoustic patterns of vocal expression. This includes the capability to learn vocal patterns by imitation, as well as the production of new patterns by invention. These abilities are essential in the uniquely human inventions of language and music. Among the primates, only humans have gained direct cortical control over the voice, which is a prerequisite for singing. Neuroanatomical studies have indicated that nonhuman primates lack the direct connection between the primary motor cortex and the nucleus ambiguus (i.e., the site of the laryngeal motoneurons) that humans have (Jürgens, 1976; Kuypers, 1958).

Comparative research. Results from neurophysiological research that indicate that there is phylogenetic continuity of vocal expression have encouraged some researchers to embark on comparative studies of vocal expression. Although biologists and ethologists have tended to shy away from using words such as *emotion* in connection with animal behavior (Plutchik, 1994, chap. 10; Scherer, 1985), a case could be made that most animal vocalizations involve motivational states that are closely related to emotions (Goodall, 1986; Hauser, 2000; Marler, 1977; Ploog, 1986; Richman, 1987; Scherer, 1985; Snowdon, 2003). The states usually have to be inferred from the specific situations in which the vocalizations occurred. "In most of the circumstances in which animal signaling occurs, one detects urgent and demanding functions to be served, often involving emergencies for survival or procreation" (Marler, 1977, p. 54). There is little systematic work on vocal expression in animals, but several studies have indicated a close correspondence between the acoustic characteristics of animal vocalizations and specific affective situations (for reviews, see Plutchik, 1994, chap. 9–10; Scherer, 1985; Snowdon, 2003). For instance, Ploog (1981, 1986) discovered a limited number of vocal expression categories in squirrel monkeys. These categories were related to important events in the monkeys' lives and included warning calls (*alarm peeps*), threat calls (*groaning*), desire for social contact calls (*isolation peeps*), and companionship calls (*cackling*).

Given phylogenetic continuity of vocal expression and cross-species similarity in the kinds of situations that generate vocal expression, it is interesting to ask whether there is any evidence of cross-species universality of vocal expression. Limited evidence of this kind has indeed been found (Scherer, 1985; Snowdon, 2003). For instance, E. S. Morton (1977) noted that "birds and mammals use harsh, relatively low-frequency sounds when hostile, and higher-frequency, more pure tonelike sounds when frightened, appeasing, or approaching in a friendly manner" (p. 855; see also Ohala, 1983). Another general principle, proposed by Jürgens (1979), is that increasing aversiveness of primate vocal calls is correlated with pitch, total pitch range, and irregularity of pitch contours. These features have also been associated with negative emotion in human vocal expression (Davitz, 1964b; Scherer, 1986).

Physiological differentiation. In animal studies, descriptions of vocal characteristics and emotional states are necessarily imprecise (Scherer, 1985), making direct comparisons difficult. However, at the least these data suggest that there are some systematic relationships among acoustic measures and emotions. An important question is how such relationships and examples of cross-species universality may be explained. According to Spencer's law, there should be common physiological principles. In fact, physiological variables determine to a large extent the nature of phonation and resonance in vocal expression (Scherer, 1989), and there may be *some* reliable differentiation of physiological patterns for discrete emotions (Cacioppo et al., 2000, p. 180).

It might be assumed that distinct physiological patterns reflect environmental demands on behavior: "Behaviors such as withdrawal, expulsion, fighting, fleeing, and nurturing each make different physiological demands. A most important function of emotion is to create the optimal physiological milieu to support the particular behavior that is called forth" (Levenson, 1994, p. 124). This process involves the central, somatic, and autonomic nervous systems. For example, fear is associated with a motivation to flee and brings about sympathetic arousal consistent with this action involving increased cardiovascular activation, greater oxygen exchange, and increased glucose availability (Mayne, 2001). Many physiological changes influence aspects of voice production, such as respiration, vocal fold vibration, and articulation, in well-differentiated ways. For instance, anger yields increased tension in the laryngeal musculature coupled with increased subglottal air pressure. This changes the production of sound at the glottis and hence changes the timbre of the voice (Johnstone & Scherer, 2000). In other words, depending on the specific physiological state, one may expect to find specific acoustic features in the voice.

This general principle underlies Scherer's (1985) component process theory of emotion, which is the most promising attempt to formulate a stringent theory along the lines of Spencer's law. Using this theory, Scherer (1985) made detailed predictions about the patterns of acoustic cues (bits of information) associated with different emotions. The predictions were based on the idea that emotions involve sequential cognitive appraisals, or *stimulus evaluation checks* (SECs), of stimulus features such as novelty, intrinsic pleasantness, goal significance, coping potential, and norm or self compatibility (for further elaboration of appraisal dimensions, see Scherer, 2001). The outcome of each SEC is assumed to have a specific effect on the somatic nervous system, which in turn affects the musculature associated with voice production. In addition, each SEC outcome is assumed to affect various aspects of the autonomous nervous system (e.g., mucous and saliva production) in ways that strongly influence voice production. Scherer (1985) did not favor the basic emotions approach, although he offered predictions for acoustic cues associated with anger, disgust, fear, happiness, and sadness—"five major types of emotional states that can be expected to occur frequently in the daily life of many organisms, both animal and human" (p. 227). Later in this review, we provide a comparison of empirical findings from vocal expression and music performance with Scherer's (1986) revised predictions.

Although human vocal expression of emotion is based on phylogenetically old parts of the brain that are in some respects similar to those of nonhuman primates, what is characteristic of humans is that they have much greater voluntary control over their vocaliza-

tion (Jürgens, 2002). Therefore, an important distinction must be made between so-called push and pull effects in the determinants of vocal expression (Scherer, 1989). Push effects involve various physiological processes, such as respiration and muscle tension, that are naturally influenced by emotional response. Pull effects, on the other hand, involve external conditions, such as social norms, that may lead to strategic posing of emotional expression for manipulative purposes (e.g., Krebs & Dawkins, 1984). Vocal expression of emotions typically involves a combination of push and pull effects, and it is generally assumed that posed expression tends to be modeled on the basis of natural expression (Davitz, 1964c, p. 16; Owren & Bachorowski, 2001, p. 175; Scherer, 1985, p. 210). However, the precise extent to which posed expression is similar to natural expression is a question that requires further research.

Vocal Expression and Music Performance: Are They Related?

It is a recurrent notion that music is a means of emotional expression (Budd, 1985; S. Davies, 2001; Gabrielsson & Juslin, 2003). Indeed, music has been defined as "one of the fine arts which is concerned with the combination of sounds with a view to beauty of form and the expression of emotion" (D. Watson, 1991, p. 8). It has been difficult to explain why music is expressive of emotions, but one possibility is that music is reminiscent of vocal expression of emotions.

Previous perspectives. The notion that there is a close relationship between music and the human voice has a long history (Helmholtz, 1863/1954; Kivy, 1980; Rousseau, 1761/1986; Scherer, 1995; Spencer, 1857; Sundberg, 1982). Helmholtz (1863/1954)—one of the pioneers of music psychology—noted that "an endeavor to imitate the involuntary modulations of the voice, and make its recitation richer and more expressive, may therefore possibly have led our ancestors to the discovery of the first means of musical expression" (p. 371). This impression is reinforced by the voicelike character of most musical instruments: "There are in the music of the violin . . . accents so closely akin to those of certain contralto voices that one has the illusion that a singer has taken her place amid the orchestra" (Marcel Proust, as cited in D. Watson, 1991, p. 236). Richard Wagner, the famous composer, noted that "the oldest, truest, most beautiful organ of music, the origin to which alone our music owes its being, is the human voice" (as cited in D. Watson, 1991, p. 2). Indeed, Stendhal commented that "no musical instrument is satisfactory except in so far as it approximates to the sound of the human voice" (as cited in D. Watson, 1991, p. 309). Many performers of blues music have been attracted to the vocal qualities of the slide guitar (Erlewine, Bogdanov, Woodstra, & Koda, 1996). Similarly, people often refer to the musical aspects of speech (e.g., Besson & Friederici, 1998; Fónagy & Magdics, 1963), particularly in the context of infant-directed speech, where mothers use changes in duration, pitch, loudness, and timbre to regulate the infant's level of arousal (M. Papoušek, 1996).

The hypothesis that vocal expression and music share a number of expressive features might appear trivial in the light of all the arguments by different authors. However, these comments are primarily anecdotal or speculative in nature. Indeed, many authors have disputed this hypothesis. S. Davies (2001) observed that

it has been suggested that expressive instrumental music recalls the tones and intonations with which emotions are given vocal expression (Kivy, 1989), but this . . . is dubious. It is true that blues guitar and jazz saxophone sometimes imitate singing styles, and that singing styles sometimes recall the sobs, wails, whoops, and yells that go with ordinary occasions of expressiveness. For the general run of cases, though, music does not sound very like the noises made by people gripped by emotion. (p. 31)

(See also Budd, 1985, p. 148; Levman, 2000, p. 194.) Thus, awaiting relevant data, it has been uncertain whether Spencer's law can provide an account of music's expressiveness.

Boundary conditions of Spencer's law. It does actually seem unlikely that Spencer's law can explain all of music's expressiveness. For instance, there are several aspects of musical form (e.g., harmonic progression) that have no counterpart in vocal expression but that nonetheless contribute to music's expressiveness (e.g., Gabrielsson & Juslin, 2003). Consequently, Spencer's law cannot be the whole story of music's expressiveness. In fact, there are many sources of emotion in relation to music (e.g., Sloboda & Juslin, 2001) including musical expectancy (Meyer, 1956), arbitrary association (J. B. Davies, 1978), and iconic signification—that is, structural similarity between musical and extramusical features (Langer, 1951). Only the last of these sources corresponds to Spencer's law. Yet, we argue that Spencer's law should be part of any satisfactory account of music's expressiveness. For the hypothesis to have explanatory power, however, it must be constrained. What is required, we propose, is specification of the boundary conditions of the hypothesis.

We argue that the hypothesis that there is an iconic similarity between vocal expression of emotion and musical expression of emotion applies only to certain acoustic features—primarily those features of the music that the performer can control (more or less freely) during his or her performance such as tempo, loudness, and timbre. However, the hypothesis does not apply to such features of a piece of music that are usually indicated in the notation of the piece (e.g., harmony, tonality, melodic progression), because these features reflect to a larger extent characteristics of music as a human art form that follows its own intrinsic rules and that varies from one culture to another (Carterette & Kendall, 1999; Juslin, 1997c). Neuropsychological research indicates that certain aspects of music (e.g., timbre) share the same neural resources as speech, whereas others (e.g., tonality) draw on resources that are unique to music (Patel & Peretz, 1997; see also Peretz, 2002). Thus, we argue that musicians communicate emotions to listeners via their performances of music by using emotion-specific patterns of acoustic cues derived from vocal expression of emotion (Juslin, 1998). The extent to which Spencer's law can offer an explanation of music's expressiveness is directly proportional to the relative contribution of performance variables to the listener's perception of emotions in music. Because performance variables include such perceptually salient features as speed and loudness, this contribution is likely to be large.

It is well-known that the same sentence may be pronounced in a large number of different ways, and that the way in which it is pronounced may convey the speaker's state of emotion. In principle, one can separate the verbal message from its acoustic realization in speech. Similarly, the same piece of music can be played in a number of different ways, and the way in which it is played may convey specific emotions to listeners. In principle, one can sepa-

rate the structure of the piece, as notated, from its acoustic realization in performance. Therefore, to obtain possible similarities, how speakers and musicians express emotions through the ways in which they convey verbal and musical contents should be explored (i.e., "It's not what you say, it's how you say it").

The origins of the relationship. If musical expression of emotion should turn out to resemble vocal expression of emotion, how did musical expression come to resemble vocal expression in the first place? The origins of music are, unfortunately, forever lost in the history of our ancestors (but for a survey of theories, see various contributions in Wallin, Merker, & Brown, 2000). However, it is apparent that music accompanies many important human activities, and this is especially true of so-called preliterate cultures (e.g., Becker, 2001; Gregory, 1997). It is possible to speculate that the origin of music is to be found in various cultural activities of the distant past, when the demarcation between vocal expression and music was not as clear as it is today. Vocal expression of discrete emotions such as happiness, sadness, anger, and love probably became gradually meshed with vocal music that accompanied related cultural activities such as festivities, funerals, wars, and caregiving. A number of authors have proposed that music served to harmonize the emotions of the social group and to create cohesion: "Singing and dancing serves to draw groups together, direct the emotions of the people, and prepare them for joint action" (E. O. Wilson, 1975, p. 564). There is evidence that listeners can accurately categorize songs of different emotional types (e.g., festive, mourning, war, lullabies) that come from different cultures (Eggebrecht, 1983) and that there are similarities in certain acoustic characteristics used in such songs; for instance, mourning songs typically have slow tempo, low sound level, and soft timbre, whereas festive songs have fast tempo, high sound level, and bright timbre (Eibl-Eibesfeldt, 1989, p. 695). Thus, it is reasonable to hypothesize that music developed from a means of emotion sharing and communication to an art form in its own right (e.g., Juslin, 2001b; Levman, 2000, p. 203; Storr, 1992, p. 23; Zucker, 1946, p. 85).

Theoretical Predictions

In the foregoing, we outlined an evolutionary perspective according to which music performers are able to communicate basic emotions to listeners by using a nonverbal code that derives from vocal expression of emotion. We hypothesized that vocal expression is an evolved mechanism based on innate, fairly stable, and universal affect programs that develop early and are fine-tuned by prenatal experiences (Mastropieri & Turkewitz, 1999; Verny & Kelly, 1981). We made the following five predictions on the basis of this evolutionary approach. First, we predicted that communication of basic emotions would be accurate in both vocal and musical expression. Second, we predicted that there would be cross-cultural accuracy of communication of basic emotions in both channels, as long as certain acoustic features are involved (speed, loudness, timbre). Third, we predicted that the ability to recognize basic emotions in vocal and musical expression develops early in life. Fourth, we predicted that the same patterns of acoustic cues are used to communicate basic emotions in both channels. Finally, we predicted that the patterns of cues would be consistent with Scherer's (1986) physiologically based predictions.

These five predictions are addressed in the following empirical review.

Definitions and Method of the Review

Basic Issues and Terminology in Nonverbal Communication

Vocal expression and music performance arguably belong to the general class of nonverbal communication behavior. Fundamental issues concerning nonverbal communication include (a) the content (What is communicated?), (b) the accuracy (How well is it communicated?), and (c) the code usage (How is it communicated?). Before addressing these questions, one should first make sure that communication has occurred. Communication implies (a) a socially shared code, (b) an encoder who intends to express something particular via that code, and (c) a decoder who responds systematically to that code (e.g., Shannon & Weaver, 1949; Wiener, Devoe, Rubinow, & Geller, 1972). True communication has taken place only if the encoder's expressive intention has become mutually known to the encoder and the decoder (e.g., Ekman & Friesen, 1969). We do not exclude that information may be unwittingly transmitted from one person to another, but this would not count as communication according to the present definition (for a different view, see Buck, 1984, pp. 4–5).

An important aspect of the communicative process is the coding of the nonverbal signs (the manner in which information is transmitted through the signal). According to Ekman and Friesen (1969), the nature of the coding can be described by three dimensions: discrete versus continuous, probabilistic versus invariant, and iconic versus arbitrary. Nonverbal signals are typically coded continuously, probabilistically, and iconically. To illustrate, (a) the loudness of the voice changes continuously (rather than discretely); (b) increases in loudness frequently (but not always) signify anger; and (c) the loudness is iconically (rather than arbitrarily) related to the intensity of the felt anger (e.g., the loudness increases when the felt intensity of the anger increases; Juslin & Laukka, 2001, Figure 4).

The Standard Content Paradigm

Studies of vocal expression and studies of music performance have typically been carried out separately from each other. However, one could argue that the two domains share a number of important characteristics. First, both domains are concerned with a channel that uses patterns of pitch, loudness, and duration to communicate emotions (the content). Second, both domains have investigated the same questions (How accurate is the communication?, What is the nature of the code?). Third, both domains have used similar methods (decoding experiments, acoustic analyses). Hence, both domains have confronted many of the same problems (see the Discussion section).

In a typical study of communication of emotions in vocal expression or music performance, the encoder (speaker/performer in vocal expression/ music performance, respectively) is presented with material to be spoken/ performed. The material usually consists of brief sentences or melodies. Each sentence/melody is to be spoken/performed while expressing different emotions prechosen by the experimenter. The emotion portrayals are recorded and used in listening tests to study whether listeners can decode the expressed emotions. Each portrayal is analyzed to see what acoustic cues are used in the communicative process. The assumption is that, because the verbal/musical material remains the same in different portrayals, whatever effects that appear in listeners' judgments or acoustic measures should primarily be the result of the encoder's expressive intention. This procedure, often referred to as *the standard content paradigm* (Davitz, 1964b), is not without its problems, but we temporarily postpone our critique until the Discussion section.

Table 1
Classification of Emotion Words Used by Different Authors Into Emotion Categories

Emotion category	Emotion words used by authors
Anger	Aggressive, aggressive–excitable, aggressiveness, anger, anger–hate–rage, angry, ärger, ärgerlich, cold anger, colère, collera, destruction, frustration, fury, hate, hot anger, irritated, rage, repressed anger, wut
Fear	Afraid, angst, ängstlich, anxiety, anxious, fear, fearful, fear of death, fear–pain, fear–terror–horror, frightened, nervousness, panic, paura, peur, protection, scared, schreck, terror, worry
Happiness	Cheerfulness, elation, enjoyment, freude, freudig, gioia, glad, glad–quiet, happiness, happy, happy–calm, happy–excited, joie, joy, laughter–glee–merriment, serene–joyful
Sadness	Crying despair, depressed–sad, depression, despair, gloomy–tired, grief, quiet sorrow, sad, sad–depressed, sadness, sadness–grief–crying, sorrow, trauer, traurig, traurigkeit, tristesse, tristezza
Love–tenderness	Affection, liebe, love, love–comfort, loving, soft–tender, tender, tenderness, tender passion, tenerezza, zärtlichkeit

Criteria for Inclusion of Studies

We used two criteria for inclusion of studies in the present review. First, we included only studies focusing on nonverbal aspects of speech or performance-related aspects of music. This is in accordance with the boundary conditions of the hypothesis discussed above. Second, we included only studies that investigated the communication of discrete emotions (e.g., sadness). Hence, studies that focused on emotional arousal in general (e.g., Murray, Baber, & South, 1996) or on emotion dimensions (e.g., Laukka, Juslin, & Bresin, 2003) were not included in the review. Similarly, studies that used the standard paradigm but that did not use explicitly defined emotions (e.g., Cosmides, 1983) or that used only positive versus negative affect (e.g., Fulcher, 1991) were not included. Such studies do not allow for the relevant comparisons with studies of music performance, which have almost exclusively studied discrete emotions.

Search Strategy

Emotion in vocal expression and music performance is a multidisciplinary field of research. The majority of studies have been conducted by psychologists, but contributions also come from, for instance, acoustics, speech science, linguistics, medicine, engineering, computer science, and musicology. Publications are scattered among so many sources that even many review articles have not surveyed more than a subset of the literature. To ensure that this review was as complete as possible, we searched for relevant investigations by using a variety of sources. More specifically, the studies included in the present review were gathered using the following Internet-based scientific databases: PsycINFO, MEDLINE, Linguistics and Language Behavior, Ingenta, and RILM Abstracts of Music Literature. Whenever possible, the year limits were set at articles published since 1900. The following words, in various combinations and truncations, were used in the literature search: *emotion*, *affective*, *vocal*, *voice*, *speech*, *prosody*, *paralanguage*, *music*, *music performance*, and *expression*. The goal was to include all English language publications in peer-reviewed journals. We have also included additional studies located via informal sources, including studies reported in conference proceedings, in other languages, and in unpublished doctoral dissertations that we were able to locate. It should be noted that the majority of studies in both domains correspond to the selection criteria above. We located 104 studies of vocal expression and 41 studies of music performance in our literature search, which was completed in June 2002.

Emotional States and Terminology

We review the findings in terms of five general categories of emotion: anger, fear, happiness, sadness, and love–tenderness, primarily because these are the only five emotion categories for which there is enough evidence in both vocal expression and music performance. They roughly correspond to the basic emotions described earlier.[4] These five categories represent a reasonable point of departure because all of them comprise what are regarded as typical emotions by lay people (Shaver et al., 1987; Shields, 1984). There is also evidence that these emotions closely correspond to the first emotion terms children learn to use (e.g., Camras & Allison, 1985) and that they serve as basic-level categories in cognitive representations of emotions (e.g., Shaver et al., 1987). Their role in musical expression of emotions is highlighted by questionnaire research (Lindström, Juslin, Bresin, & Williamon, 2003) in which 135 music students were asked what emotions can be expressed in music. *Happiness*, *sadness*, *fear*, *love*, and *anger* were among the 10 most highly rated words of a list of 38 words containing both basic and complex emotions.

An important question concerns the exact words used to denote the emotions. A number of different words have been used in the literature, and there is little agreement so far regarding the organization of the emotion lexicon (Plutchik, 1994, p. 45). Therefore, it is not clear how words such as *happiness* and *joy* should be distinguished. The most prudent approach to take is to treat different but closely related emotion words (e.g., *sorrow*, *grief*, *sadness*) as belonging to the same emotion family (e.g., the sadness family; Ekman, 1992). Table 1 shows how the emotion words used in the present studies have been categorized in this review (for some empirical support, see the analyses of emotion words presented by Johnson-Laird & Oatley, 1989; Shaver at al., 1987).

Studies of Vocal Expression: Overview

Darwin (1872/1998) discussed both vocal and facial expression of emotions in his treatise. In recent years, however, facial expression has received far more empirical research than vocal expression. There are a number of reasons for this, such as the problems associated with the recording and analysis of speech sounds (Scherer, 1982). The consequence is that the code used in facial expression of emotion is better understood than the code used in vocal expression. This is unfortunate, however, because recent studies using self-reports have revealed that, if anything,

---

[4] Most theorists distinguish between *passionate love* (eroticism) and *companionate love* (tenderness; Hatfield & Rapson, 2000, p. 660), of which the latter corresponds to our love–tenderness category. Some researchers suggest that all kinds of love originally derived from this emotional state, which is associated with infant–caregiver attachment (e.g., Eibl-Eibesfeldt, 1989, chap. 4; Oatley & Jenkins, 1996; p. 287; Panksepp, 1998, chap. 13).

vocal expressions may be even more important predictors of emotions than facial expressions in everyday life (Planalp, 1998). Fortunately, the field of vocal expression of emotions has recently seen renewed interest (Cowie, Douglas-Cowie, & Schröder, 2000; Cowie et al., 2001; Johnstone & Scherer, 2000). Thirty-two studies were published in the 1990s, and already 19 studies have been published between January 2000 and June 2002.

Table 2 provides a summary of 104 studies of vocal expression included in this review in terms of authors, publication year, emotions studied, method used (e.g., portrayal, manipulated portrayal, induction, natural speech sample, synthesis), language, acoustic cues analyzed (where applicable), and verbal material. Thirty-nine studies presented data that permitted us to include them in a meta-analysis of communication accuracy (detailed below). The majority of studies (58%) used English-speaking encoders, although as many as 18 different languages, plus nonsense utterances, are represented in the studies reviewed. Twelve studies (12%) can be characterized as more or less cross-cultural in that they included analyses of encoders or decoders from more than one nation. The verbal material features series of numbers, letters of the alphabet, nonsense syllables, or regular speech material (e.g., words, sentences, paragraphs). The number of emotions included ranges from 1 to 15 ($M = 5.89$). Ninety studies (87%) used emotion portrayals by actors, 13 studies (13%) used manipulations of portrayals (e.g., filtering, masking, reversal), 7 studies (7%) used mood induction procedures, and 12 studies (12%) used natural speech samples. The latter comes mainly from studies of fear expressions in aviation accidents. Twenty-one studies (20%) used sound synthesis, or *copy synthesis*.[5] Seventy-seven studies (74%) reported acoustic data, of which 6 studies used listeners' ratings of cues rather than acoustic measurements.

Studies of Music Performance: Overview

Studies of music performance have been conducted for more than 100 years (for reviews, see Gabrielsson, 1999; Palmer, 1997). However, these studies have almost exclusively focused on structural aspects of performance such as marking of the phrase structure, whereas emotion has been ignored. Those studies that have been concerned with emotion in music, on the other hand, have almost exclusively focused on expressive aspects of musical composition such as pitch or mode (e.g., Gabrielsson & Juslin, 2003), whereas they have ignored aspects of specific performances. That performance aspects of emotional expression did not gain attention much earlier is strange considering that one of the great pioneers in music psychology, Carl E. Seashore, made detailed proposals about such studies in the 1920s (Seashore, 1927). Seashore (1947) later suggested that music researchers could use the same paradigm that had been used in vocal expression (i.e., the standard content paradigm) to investigate how performers express emotions. However, Seashore's (1947) plea went unheard, and he did not publish any study of that kind himself. After slow initial progress, there was an increase of studies in the 1990s (23 studies published). This seems to continue into the 2000s (10 studies published 2000–2002). The increase is perhaps a result of the increased availability of software for digital analysis of acoustic cues, but it may also reflect a renaissance for research on musical emotion (Juslin & Sloboda, 2001). Figure 1 illustrates the timeliness of this review in terms of the studies available for a comparison of the two domains.

Table 3 provides a summary of the 41 studies of emotional expression in music performance included in the review in terms of authors, publication year, emotions studied, method used (e.g., portrayal, manipulated portrayal, synthesis), instrument used, number and nationality of performers and listeners, acoustic cues analyzed (where applicable), and musical material. Twelve studies (29%) provided data that permitted us to include them in a meta-analysis of communication accuracy. These studies covered a wide range of musical styles, including classical music, folk music, Indian ragas, jazz, pop, rock, children's songs, and free improvisations.

The most common musical style was classical music (17 studies, 41%). Most studies relied on the standard paradigm used in studies of vocal expression of emotions. The number of emotions studied ranges from 3 to 9 ($M = 4.98$), and emotions typically included happiness, sadness, anger, fear, and tenderness. Twelve musical instruments were included. The most frequently studied instrument was singing voice (19 studies), followed by guitar (7), piano (6), synthesizer (4), violin (3), flute (2), saxophone (2), drums (1), sitar (1), timpani (1), trumpet (1), xylophone (1), and sentograph (1—a sentograph is an electronic device for recording patterns of finger pressure over time—see Clynes, 1977). At least 12 different nationalities are represented in the studies (Fónagy & Magdics, 1963, did not state the nationalities clearly), with Sweden being most strongly represented (39%), followed by Japan (12%) and the United States (12%). Most of the studies analyzed professional musicians (but see Juslin & Laukka, 2000), and the performances were usually monophonic to facilitate measurement of acoustic parameters (for an exception, see Dry & Gabrielsson, 1997). (Monophonic melody is probably one of the earliest forms of music, Wolfe, 2002.) A few studies (15%) investigated what means listeners use to decode emotions by means of synthesized performances. Eighty-five percent of the studies reported data on acoustic cues; of these studies, 5 used listeners' ratings of cues rather than acoustic measurements.

Results

Decoding Accuracy

Studies of vocal expression and music performance have converged on the conclusion that encoders can communicate basic emotions to decoders with above-chance accuracy, at least for the five emotion categories considered here. To examine these data closer, we conducted a meta-analysis of decoding accuracy. Included in this analysis were all studies that presented (or allowed computation of) forced-choice decoding data relative to some independent criterion of encoding intention. Thirty-nine studies of vocal expression and 12 studies of music performance met this criterion, featuring a total of 73 decoding experiments, 60 for vocal expression, 13 for music performance.

One problem in comparing accuracy scores from different studies is that they use different numbers of response alternatives in the decoding task. Rosenthal and Rubin's (1989) effect size index for one-sample, multiple–choice-type data, pi (π), allows researchers to transform accuracy scores involving any number of response alternatives to a standard scale of dichotomous choice, on which .50 is the null value and 1.00 corresponds to 100% correct decoding. Ideally, an index of decoding accuracy should also take into account the response bias in the decoder's judgments (Wagner, 1993). However, this requires that results be presented in terms of a confusion matrix, which very few studies have done. Therefore, we summarize the data simply in terms of Rosenthal and Rubin's pi index.

Summary statistics. Table 4 summarizes the main findings from the meta-analysis in terms of summary statistics (i.e., unweighted mean, weighted mean, median, standard deviation,

---

[5] Copy synthesis refers to copying acoustic features from real emotion portrayals and using them to resynthesize new portrayals. This method makes it possible to manipulate certain cues of an emotion portrayal while leaving other cues intact (e.g., Juslin & Madison, 1999; Ladd et al., 1985; Schröder, 2001).

Table 2
Summary of Studies on Vocal Expression of Emotion Included in the Review

Study	Emotions studied (in terms used by authors)	Method	Speakers/listeners		Acoustic cues analyzed[a]	Verbal material
			N	Language		
1. Abelin & Allwood (2000)	Anger, disgust, dominance, fear, happiness, sadness, surprise, shyness	P	1/93	Swe/Eng (12), Swe/Fin (23), Swe/Spa (23), Swe/Swe (35)	SR, pauses, F0, Int	Brief sentence
2. Albas et al. (1976)	Anger, happiness, love, sadness	P, M	12/80	Eng (6)/Eng (20), Cre (20), Cre (6)/Eng (20), Cre (20)		"Any two sentences that come to mind"
3. Al-Watban (1998)	Anger, fear, happiness, neutral, sadness	P	4/14	Ara	SR, F0, Int	Brief sentence
			4/15	Eng		
4. Anolli & Ciceri (1997)	Collera [anger], disprezzo [disgust], gioia [happiness], paura [fear], tenerezza [tenderness], tristezza [sadness]	P, M	2/100	Ita	SR, pauses, F0, Int, Spectr	Brief sentence
5. Apple & Hecht (1982)	Anger, happiness, sadness, surprise	P, M	43/48	Eng		Brief sentence
6. Banse & Scherer (1996)	Anxiety, boredom, cold anger, contempt, despair, disgust, elation, happiness, hot anger, interest, panic, pride, sadness, shame	P	12/12	Non/Ger	SR, F0, Int, Spectr	Brief sentence
7. Baroni et al. (1997)	Anger, happiness, sadness	P	3/42	Ita	SR, F0, Int	Brief sentence
8. Baroni & Finarelli (1994)	Aggressive, depressed–sad, serene–joyful	P	3/0	Ita	SR, Int	Brief sentence
9. Bergmann et al. (1988)	Ärgerlich [angry], ängstlich [afraid], entgegenkommend [accomodating], freudig [glad], gelangweilt [bored], nachdrücklich [insistent], traurig [sad], verächtlich [scornful], vorwurfsvoll [reproachful]	S	0/88	Ger	SR, F0, F0 contour, Int, jitter	Brief sentence
10. Bonebright (1996)	Anger, fear, happiness, neutral, sadness	P	6/0	Eng	SR, F0, Int	Long paragraph
11. Bonebright et al. (1996)	Anger, fear, happiness, neutral, sadness (same portrayals as in Study 10)	P	6/104	Eng		Long paragraph
12. Bonner (1943)	Fear	P	3/0	Eng	SR, F0, pauses	Brief sentence
13. Breitenstein et al. (2001)	Angry, frightened, happy, neutral, sad	P, S	1/65	Ger/Ger (35), Ger/Eng (30)	SR, F0	Brief sentence
14. Breznitz (1992)	Angry, happy, sad, neutral (post hoc classification)	N	11/0	Eng	F0	Responses from interviews
15. Brighetti et al. (1980)	Anger, contempt, disgust, fear, happiness, sadness, surprise	P	6/34	Ita		Series of numbers
16. Burkhardt (2001)	Boredom, cold anger, crying despair, fear, happiness, hot anger, joy, quiet sorrow	S	0/72	Ger	SR, F0, Spectr, formants, Art., glottal waveform	Brief sentence
		P	10/30	Ger		
17. Burns & Beier (1973)	Angry, anxious, happy, indifferent, sad, seductive	P	30/21	Eng		Brief sentence
18. Cahn (1990)	Angry, disgusted, glad, sad, scared, surprised	S	0/28	Eng	SR, pauses, F0, Spectr, Art., glottal waveform	Brief sentence
19. Carlson et al. (1992)	Angry, happy, neutral, sad	P, S	1/18	Swe	SR, F0, F0 contour	Brief sentence
20. Chung (2000)	Joy, sadness	N	1/30	Kor/Kor (10), Kor/Eng (10), Kor/Fre (10)	SR, F0, F0 contour, Spectr, jitter, shimmer	Responses from TV interviews
		S	5/22	Eng		
		P	0/20	Eng		
21. Costanzo et al. (1969)	Anger, contempt, grief, indifference, love	P	23/44	Eng	(SR, F0, Int)	Brief paragraph

Table 2 (*continued*)

Study	Emotions studied (in terms used by authors)	Method	Speakers/listeners N	Speakers/listeners Language	Acoustic cues analyzed[a]	Verbal material
22. Cummings & Clements (1995)	Angry (+ 10 nonemotion terms)	P	2/0	Eng	Glottal waveform	Word
23. Davitz (1964a)	Admiration, affection, amusement, anger, boredom, cheerfulness, despair, disgust, dislike, fear, impatience, joy, satisfaction, surprise	P	7/20	Eng		Brief paragraph
24. Davitz (1964b)	Affection, anger, boredom, cheerfulness, impatience, joy, sadness, satisfaction	P	5/5	Eng	(SR, F0, F0 contour, Int, Art., rhythm, timbre)	Brief paragraph
25. Davitz & Davitz (1959)	Anger, fear, happiness, jealousy, love, nervousness, pride, sadness, satisfaction, sympathy	P	8/30	Eng		Letters of the alphabet
26. Dusenbury & Knower (1939)	Amazement–astonishment–surprise, anger–hate–rage, determination–stubbornness–firmness, doubt–hesitation–questioning, fear–terror–horror, laughter–glee–merriment, pity–sympathy–helpfulness, religious love–reverence–awe, sadness–grief–crying, sneering–contempt–scorn, torture–great pain–suffering	P	8/457	Eng		Letters of the alphabet
27. Eldred & Price (1958)	Anger, anxiety, depression	N	1/4	Eng	(SR, pauses, F0, Int)	Speech from psychotherapy sessions
28. Fairbanks & Hoaglin (1941)	Anger, contempt, fear, grief, indifference	P	6/0	Eng	SR, pauses	Brief paragraph
29. Fairbanks & Provonost (1938)	Anger, contempt, fear, grief, indifference (same portrayals as in Study 28)	P	6/64	Eng	F0	Brief paragraph
30. Fairbanks & Provonost (1939)	Anger, contempt, fear, grief, indifference (same portrayals as in Study 28)	P	6/64	Eng	F0, F0 contour, pauses	Brief paragraph
31. Fenster et al. (1977)	Anger, contentment, fear, happiness, love, sadness	P	5/30	Eng	F0, F0 contour	Brief sentence
32. Fónagy (1978)	Anger, coquetry, disdain, fear, joy, longing, repressed anger, reproach, sadness, tenderness	P, M	1/58	Fre	F0, F0 contour	Brief sentence
33. Fónagy & Magdics (1963)	Anger, complaint, coquetry, fear, joy, longing, sarcasm, scorn, surprise, tenderness	P	—/0	Hun, Fre, Ger, Eng	SR, F0, F0 contour, Int, timbre	Brief sentence
34. Frick (1986)	Anger (frustration, threat), disgust	P	1/37	Eng	F0	Brief sentence
35. Friend & Farrar (1994)	Angry, happy, neutral	P, M	1/99	Eng	F0, Int, Spectr	Brief sentence
36. Gårding & Abramson (1965)	Anger, surprise, neutral (+ 2 nonemotion terms)	P; S	5/5; 0/11	Eng	F0, F0 contour	Brief sentence, digits
37. Gérard & Clément (1998)	Happiness, irony, sadness (+ 2 nonemotion terms)	P	3/0; 1/20	Fre	SR, F0, F0 contour	Brief sentence
38. Gobl & Ní Chasaide (2000)	Afraid/unafraid, bored/interested, content/angry, friendly/hostile, relaxed/stressed, sad/happy, timid/confident	S	0/8	Swe	Int, jitter, formants, Spectr, glottal waveform	Brief sentence
39. Graham et al. (2001)	Anger, depression, fear, hate, joy, nervousness, neutral, sadness	P	4/177	Eng/Eng (85), Eng/Jap (54), Eng/Spa (38)	F0	Long paragraph
40. Greasley et al. (2000)	Anger, disgust, fear, happiness, sadness	N	32/158	Eng		Speech from TV and radio

(*table continues*)

Table 2 (*continued*)

Study	Emotions studied (in terms used by authors)	Method	N	Language	Acoustic cues analyzed[a]	Verbal material
				Speakers/listeners		
41. Guidetti (1991)	Colère [anger], joie [joy], neutre [neutral], peur [fear], tristesse [sadness]	P	4/50	Non/Fre		Brief sentence
42. Havrdova & Moravek (1979)	Anger, anxiety, joy	I	6/0	Cze	SR, F0, Int, Spectr	Brief sentence
43. Höffe (1960)	Ärger [anger], enttäuschung [disappointment], erleichterung [relief], freude [joy], schreck [fear], schmerz [pain], trost [comfort], trotz [defiance], wohlbehagen [satisfaction], zweifel [doubt]	P	4/0	Ger	F0, Int	Word
44. House (1990)	Angry, happy, neutral, sad	P, M, S	1/11	Swe	F0, F0 contour, Int	Brief sentence
45. Huttar (1968)	Afraid/bold, angry/pleased, sad/happy, timid/confident, unsure/sure	N	1/12	Eng	SR, F0, Int	Speech from classroom discussions
46. Iida et al. (2000)	Anger, joy, sadness	I S	2/0 0/36	Jap	SR, F0, Int	Paragraph
47. Iriondo et al. (2000)	Desire, disgust, fear, fury, joy, sadness, surprise (three levels of intensity)	P S	8/1,054 0/—	Spa	SR, pauses, F0, Int, Spectr	Paragraph
48. Jo et al. (1999)	Afraid, angry, happy, sad	P S	1/— 0/20	Kor	SR, F0	Brief sentence
49. Johnson et al. (1986)	Anger, fear, joy, sadness	P, S M	1/21 1/23	Eng		Brief sentence
50. Johnstone & Scherer (1999)	Anxious, bored, depressed, happy, irritated, neutral, tense	P	8/0	Fre	F0, Int, Spect, jitter, glottal waveform	Brief sentence, phoneme, numbers
51. Juslin & Laukka (2001)	Anger, disgust, fear, happiness, no expression, sadness (two levels of intensity)	P	8/45	Eng (4), Swe (4)/Swe (45)	SR, pauses, F0, F0 contour, Int, attack, formants, Spectr, jitter, Art.	Brief sentence
52. L. Kaiser (1962)	Cheerfulness, disgust, enthusiasm, grimness, kindness, sadness	P	8/51	Dut	SR, F0, F0 contour, Int, formants, Spectr	Vowels
53. Katz (1997)	Anger, disgust, happy–calm, happy–excited, neutral, sadness	I	100/0	Eng	SR, F0, Int	Brief sentence
54. Kienast & Sendlmeier (2000)	Anger, boredom, fear, happiness, sadness	P	10/20	Ger	Spectr, formants, Art.	Brief sentence
55. Kitahara & Tohkura (1992)	Anger, joy, neutral, sadness	P M S	1/8 1/7 0/27	Jap	SR, F0, Int	Brief sentence
56. Klasmeyer & Sendlmeier (1997)	Anger, boredom, disgust, fear, happiness, sadness	P	3/20	Ger	F0, Int, Spectr, jitter, glottal waveform	Brief sentence
57. Knower (1941)	Amazement–astonishment–surprise, anger–hate–rage, determination–stubbornness–firmness, doubt–hesitation–questioning, fear–terror–horror, laughter–glee–merriment, pity–sympathy–helpfulness, religious love–reverence–awe, sadness–grief–crying, sneering–contempt–scorn, torture–great pain–suffering	P, M	8/27	Eng		Letters of the alphabet

Table 2 (*continued*)

Study	Emotions studied (in terms used by authors)	Speakers/listeners			Acoustic cues analyzed[a]	Verbal material
		Method	N	Language		
58. Knower (1945)	Amazement–astonishment–surprise, anger–hate–rage, determination–stubbornness–firmness, doubt–hesitation–questioning, fear–terror–horror, laughter–glee–merriment, pity–sympathy–helpfulness, religious love–reverence–awe, sadness–grief–crying, sneering–contempt–scorn, torture–great pain–suffering	P	169/15	Eng		Letters of the alphabet
59. Kramer (1964)	Anger, contempt, grief, indifference, love	P, M	10/27	Eng (7)/Eng, Jap (3)/Eng		Brief paragraph
60. Kuroda et al. (1976)	Fear (emotional stress)	N	14/0	Eng		Radio communication (flight accidents)
61. Laukkanen et al. (1996)	Anger, enthusiasm, neutral, sadness, surprise	P	3/25	Non/Fin	F0, Int, glottal waveform, subglottal pressure	Brief sentence
62. Laukkanen et al. (1997)	Anger, enthusiasm, neutral, sadness, surprise (same portrayals as in Study 61)	P, S	3/10	Non/Fin	Glottal waveform, formants	Brief sentence
63. Leinonen et al. (1997)	Admiring, angry, astonished, commanding, frightened, naming, pleading, sad, satisfied, scornful	P	12/73	Fin	SR, F0, F0 contour, Int, Spectr	Word
64. Léon (1976)	Admiration [admiration], colère [anger], joie [joy], ironie [irony], neutre [neutral], peur [fear], surprise [surprise], tristesse [sadness]	P	1/20	Fre	SR, pauses, F0, F0 contour, Int	Brief sentence
65. Levin & Lord (1975)	Destruction (anger), protection (fear)	P	5/0	Eng	F0, Spectr	Word
66. Levitt (1964)	Anger, contempt, disgust, fear, joy, surprise	P	50/8	Eng		Brief paragraph
67. Lieberman (1961)	Bored, confidential, disbelief/doubt, fear, happiness, objective, pompous	P	6/20	Eng	Jitter	Brief sentence
68. Lieberman & Michaels (1962)	Boredom, confidential, disbelief, fear, happiness, pompous, question, statement	P M	6/20 3/60	Eng	F0, F0 contour, Int	Brief sentence
69. Markel et al. (1973)	Anger, depression	I	50/0	Eng	(SR, F0, Int)	Responses to the Thematic Apperception Test
70. Moriyama & Ozawa (2001)	Anger, fear, joy, sorrow	P	1/8	Jap	SR, F0, Int	Word
71. Mozziconacci (1998)	Anger, boredom, fear, indignation, joy, neutrality, sadness	P	3/10	Dut	SR, F0, F0 contour, rhythm	Brief sentence
72. Murray & Arnott (1995)	Anger, disgust, fear, grief, happiness, sadness	S S	0/52 0/35	Eng	SR, pauses, F0, F0 contour, Art., Spectr, glottal waveform	Brief sentence, paragraph
73. Novak & Vokral (1993)	Anger, joy, sadness, neutral	P	1/2	Cze	F0, Spectr	Long paragraph
74. Paeschke & Sendlmeier (2000)	Anger, boredom, fear, happiness, sadness	P	10/20	Ger	F0, Spectr	Brief sentence
75. Pell (2001)	Angry, happy, sad	P	10/10	Eng	F0, F0 contour	Brief sentence
76. Pfaff (1954)	Disgust, doubt, excitement, fear, grief, hate, joy, love, pleading, shame	P	1/304	Eng	SR, F0, F0 contour	Numbers

(*table continues*)

Table 2 (*continued*)

Study	Emotions studied (in terms used by authors)	Method	N	Language	Acoustic cues analyzed[a]	Verbal material
				Speakers/listeners		
77. Pollack et al. (1960)	Anger, approval, boredom, confidentiality, disbelief, disgust, fear, happiness, impatience, objective, pedantic, sarcasm, surprise, threat, uncertainty	P / M	4/18 / 4/28	Eng		Brief sentence
78. Protopapas & Lieberman (1997)	Terror	N / S	1/0 / 0/50	Eng	F0, jitter	Radio communication (flight accidents)
79. Roessler & Lester (1976)	Anger, affect, fear, depression	N	1/3	Eng	F0, Int, formants	Speech from psychotherapy sessions
80. Scherer (1974)	Anger, boredom, disgust, elation, fear, happiness, interest, sadness, surprise	S	0/10	Non/Am	SR, F0, F0 contour, Int	Tone sequences
81. Scherer et al. (2001)	Anger, disgust, fear, joy, sadness	P	4/428	Non/Dut (60), Non/Eng (72), Non/Fre (96), Non/Ger (70), Non/Ita (43), Non/Ind (38), Non/Spa (49)		Brief sentence
82. Scherer et al. (1991)	Anger, disgust, fear, joy, sadness	P	4/454	Non/Ger	SR, F0, Int, Spectr	Brief sentence
83. Scherer & Oshinsky (1977)	Anger, boredom, disgust, fear, happiness, sadness, surprise	S	0/48	Non/Am	SR, F0, F0 contour, Int, attack, Spectr	Tone sequences
84. Schröder (1999)	Anger, fear, joy, neutral, sadness	P / S	3/4 / 0/13	Ger		Brief sentence
85. Schröder (2000)	Admiration, boredom, contempt, disgust, elation, hot anger, relief, startle, threat, worry	P	6/20	Ger		Affect bursts
86. Sedláček & Sychra (1963)	Angst [fear], einfacher aussage [statement], feierlichkeit [solemnity], freude [joy], ironie [irony], komik [humor], liebesgefühle [love], trauer [sadness]	I	23/70	Cze	F0, F0 contour	Brief sentence
87. Simonov et al. (1975)	Anxiety, delight, fear, joy	P	57/0	Rus	F0, formants	Vowels
88. Skinner (1935)	Joy, sadness	I	19/0	Eng	F0, Int, Spectr	Word
89. Sobin & Alpert (1999)	Anger, fear, joy, sadness	I	31/12	Eng	SR, pauses, F0, Int	Brief sentence
90. Sogon (1975)	Anger, contempt, grief, indifference, love	P	4/10	Jap		Brief sentence
91. Steffen-Batóg et al. (1993)	Amazement, anger, boredom, joy, indifference, irony, sadness	P	4/70	Pol	(SR, F0, Int, voice quality)	Brief sentence
92. Stibbard (2001)	Anger, disgust, fear, happiness, sadness (same speech material as in Study 40)	N	32/—	Eng	SR, (F0, F0 contour, Int, Art., voice quality)	Speech from TV
93. Sulc (1977)	Fear (emotional stress)	N	9/0	Eng	F0	Radio communication (flight accidents)
94. Tickle (2000)	Angry, calm, fearful, happy, sad	P	6/24	Eng (3), Jap (3)/Eng (16), Jap (8)		Brief sentence, vowels
95. Tischer (1993, 1995)	Abneigung [aversion], angst [fear], ärger [anger], freude [joy], liebe [love], lust [lust], sehnsucht [longing], traurigkeit [sadness], überraschung [surprise], unsicherheit [insecurity], wut [rage], zärtlichkeit [tenderness], zufriedenheit [contentment], zuneigung [sympathy]	P	4/931	Ger	SR, Pauses, F0, Int	Brief sentence, paragraph

Table 2 (*continued*)

Study	Emotions studied (in terms used by authors)	Speakers/listeners			Acoustic cues analyzed[a]	Verbal material
		Method	N	Language		
96. Trainor et al. (2000)	Comfort, fear, love, surprise	P	23/6	Eng	SR, F0, F0 contour, rhythm	Brief sentence
97. van Bezooijen (1984)	Anger, contempt, disgust, fear, interest, joy, neutral, sadness, shame, surprise	P	8/0	Dut	SR, F0, Int, Spectr, jitter, Art.	Brief sentence
98. van Bezooijen et al. (1983)	Anger, contempt, disgust, fear, interest, joy, neutral, sadness, shame, surprise (same portrayals as in Study 97)	P	8/129	Dut/Dut (48), Dut/Tai (40), Dut/Chi (41)		Brief sentence
99. Wallbott & Scherer (1986)	Anger, joy, sadness, surprise	P M	6/11 6/10	Ger	SR, F0, Int	Brief sentence
100. Whiteside (1999a)	Cold anger, elation, happiness, hot anger, interest, neutral, sadness	P	2/0	Eng	SR, formants	Brief sentence
101. Whiteside (1999b)	Cold anger, elation, happiness, hot anger, interest, neutral, sadness (same portrayals as in Study 100)	P	2/0	Eng	F0, Int, jitter, shimmer	Brief sentence
102. Williams & Stevens (1969)	Fear	N	3/0	Eng	F0, jitter	Radio communication (flight accidents)
103. Williams & Stevens (1972)	Anger, fear, neutral, sorrow	P N	4/0 1/0	Eng Eng	SR, pauses, F0, F0 contour, Spectr, jitter, formants, Art.	Brief sentence
104. Zuckerman et al. (1975)	Anger, disgust, fear, happiness, sadness, surprise	P	40/61	Eng		Brief sentence

Note. A dash indicates that no information was provided. Types of method included portrayal (P), manipulated portrayal (M), synthesis (S), natural speech sample (N), and induction of emotion (I). Swe = Swedish; Eng = English; Fin = Finnish; Spa = Spanish; SR = speech rate; F0 = fundamental frequency; Int = voice intensity; Cre = Cree-speaking Canadian Indians; Ara = Arabic; Ita = Italian; Spectr = cues related to spectral energy distribution (e.g., high-frequency energy); Non = Nonsense utterances; Ger = German; Art. = precision of articulation; Kor = Korean; Fre = French; Hun = Hungarian; Am = American; Jap = Japanese; Cze = Czechoslovakian; Dut = Dutch; Ind = Indonesian; Rus = Russian; Pol = Polish; Tai = Taiwanese; Chi = Chinese.
[a] Acoustic cues listed within parentheses were obtained by means of listener ratings rather than acoustic measurements.

Table 3
Summary of Studies on Musical Expression of Emotion Included in the Review

Study	Emotions studied (in terms used by authors)	Method	N p/l	Nat. p/l	Instrument (n)	Acoustic cues analyzed[a]	Musical material
1. Arcos et al. (1999)	Aggressive, calm, joyful, restless, sad, tender	P, S	1/—	Spa	Saxophone	Tempo, timing, Int, Art., attack, vibrato	Jazz ballads
2. Baars & Gabrielsson (1997)	Angry, fearful, happy, no expression, sad, solemn, tender	P	1/9	Swe	Singing	Tempo, timing, Int, Art.	Folk music
3. Balkwill & Thompson (1999)	Anger, joy, peace, sadness	P	2/30	Ind/Can	Flute (1), sitar (1)	(Tempo, pitch, melodic and rhythmic complexity)	Indian ragas
4. Baroni et al. (1997)	Aggressive, sad–depressed, serene–joyful	P	3/42	Ita	Singing	Tempo, timing, Int, pitch	Opera
5. Baroni & Finarelli (1994)	Aggressive, sad–depressed, serene–joyful	P	3/0	Ita	Singing	Timing, Int	Opera
6. Behrens & Green (1993)	Angry, sad, scared	P	8/58	Am	Singing (2), trumpet (2), violin (2), timpani (2)		Improvisations
7. Bresin & Friberg (2000)	Anger, fear, happiness, no expression, sadness, solemnity, tenderness	S	0/20	Swe	Piano	Tempo, timing, Int, Art.	Children's song, classical music (romantic)
8. Bunt & Pavlicevic (2001)	Anger, fearful, happy, sad, tender	P	24/24	Eng	Var	(Tempo, timing, Int, timbre, Art., pitch)	Improvisations
9. Canazza & Orio (1999)	Active–animated, aggressive–excitable, calm–oppressive, glad–quiet, gloomy–tired, powerful–restless	P	2/17	Ita	Saxophone (1), piano (1)	Tempo, Int, Art.	Jazz
10. Dry & Gabrielsson (1997)	Angry, fearful, happy, neutral, sad, solemn, tender	P	4/20	Swe	Rock combo	(Tempo, timing, timbre, attack)	Rock music
11. Ebie (1999)	Anger, fear, happiness, sadness	P	56/3	Am	Singing	Tempo, timing, Int, Art.	Newly composed music (classical)
12. Fónagy & Magdics (1963)	Anger, complaint, coquetry, fear, joy, longing, sarcasm, scorn, surprise	P	—/—	Var	—	(Tempo, Int, timbre, Art., pitch)	Folk music, classical music, opera
13. Gabrielsson & Juslin (1996)	Angry, fearful, happy, no expression, sad, solemn, tender	P	6/37 3/56	Swe	Electric guitar Flute (1), violin (1), singing (1)	Tempo, timing, Int, Spectr, Art., pitch, attack, vibrato	Folk music, classical music, popular music
14. Gabrielsson & Lindström (1995)	Angry, happy, indifferent, soft–tender, solemn	P	4/110	Swe	Synthesizer, sentograph	Tempo, timing, Int, Art.	Folk music, popular music
15. Jansens et al. (1997)	Anger, fear, joy, neutral, sadness	P	14/25	Dut	Singing	Tempo, Int, F0, Spectr, vibrato	Classical music (lied)
16. Juslin (1993)	Anger, happiness, solemnity, tenderness, no expression	P	3/13	Swe	Electric guitar	Tempo, timing, Int, timbre, Art., attack	Folk music
17. Juslin (1997a)	Anger, fear, happiness, sadness, tenderness	P, S	1/15	Swe	Electric guitar (1), synthesizer	Tempo, timing, Int, Art., pitch	Folk music
18. Juslin (1997b)	Anger, fear, happiness, no expression, sadness	P	3/24	Swe	Electric guitar	Tempo, timing, Int, Art., attack	Jazz
19. Juslin (1997c)	Anger, fear, happiness, sadness, tenderness	P, M, S	1/12	Swe	Electric guitar (1), synthesizer	Tempo, timing, Int, Spectr, Art., attack, vibrato	Folk music
		S	0/42		Synthesizer		
20. Juslin (2000)	Anger, fear, happiness, sadness	P	3/30	Swe	Electric guitar	Tempo, Int, Spectr, Art.	Jazz, folk music

Table 3 (*continued*)

Study	Emotions studied (in terms used by authors)	Method	N p/l	Instrument (n)	Nat. p/l	Acoustic cues analyzed[a]	Musical material
21. Juslin et al. (2002)	Sadness	S	0/12	Piano	Swe	Tempo, Int, Art.	Ballad
22. Juslin & Laukka (2000)	Anger, fear, happiness, sadness	P	8/50	Electric guitar	Swe	Tempo, timing, Int, Spectr, Art.	Jazz, folk music
23. Juslin & Madison (1999)	Anger, fear, happiness, sadness	P, M	3/20	Piano	Swe	Tempo, timing, Int, Art.	Jazz, folk music
24. Konishi et al. (2000)	Anger, fear, happiness, sadness	P	10/25	Singing	Jap	Vibrato	Classical music
25. Kotlyar & Morozov (1976)	Anger, fear, joy, neutral, sorrow	P / M	11/10 / 0/11	Singing	Rus	Tempo, timing, Art., Int, attack	Classical music
26. Langeheinecke et al. (1999)	Anger, fear, joy, sadness	P	11/0	Singing	Ger	Tempo, timing, Int, Spectr, vibrato	Classical music
27. Laukka & Gabrielsson (2000)	Angry, fearful, happy, no expression, sad, solemn, tender	P	2/13	Drums	Swe	Tempo, timing, Int	Rhythm patterns (jazz, rock)
28. Madison (2000b)	Anger, fear, happiness, sadness	P, M	3/10	Piano	Swe	Timing, Int, Art. (patterns of changes)	Popular music
29. Mergl et al. (1998)	Freude [joy], trauer [sadness], wut [rage]	P	20/74	Xylophone	Ger	(Tempo, Int, Art., attack)	Improvisations
30. Metfessel (1932)	Anger, fear, grief, love	P	1/0	Singing	Am	Vibrato	Popular music
31. Morozov (1996)	Fear, hot anger, joy, neutral, sorrow	P, M	—/—	Singing	Rus	Spectr, formants	Classical music, pop, hard rock
32. Ohgushi & Hattori (1996a)	Anger, fear, joy, neutral, sorrow	P	3/0	Singing	Jap	Tempo, F0, Int, vibrato	Classical music
33. Ohgushi & Hattori (1996b)	Anger, fear, joy, neutral, sorrow	P	3/10	Singing	Jap		Classical music
34. Oura & Nakanishi (2000)	Anger, happiness, sadness	P	1/30	Piano	Jap	Tempo, Int	Classical music
35. Rapoport (1996)	Calm, excited, expressive, intermediate, neutral–soft, short, transitional–multistage, virtuoso[b]	P	—/0	Singing	Var	F0, intonation, vibrato, formants	Classical music, opera
36. Salgado (2000)	Angry, fearful, happy, neutral, sad	P	3/6	Singing	Por	Int, Spectr	Classical music (lied)
37. Scherer & Oshinsky (1977)	Anger, boredom, disgust, fear, happiness, sadness, surprise	S	0/48	Synthesizer	Am	Tempo, F0, F0 contour, Int, attack, Spectr	Classical music
38. Senju & Ohgushi (1987)	Sad (+ 9 nonemotion terms)	P	1/16	Violin	Jap		Classical music
39. Sherman (1928)	Anger–hate, fear–pain, sorrow, surprise	P, M	1/30	Singing	Am		—
40. Siegwart & Scherer (1995)	Fear of death, madness, sadness, tender passion	P	5/11	Singing	—	Int, Spectr	Opera
41. Sundberg et al. (1995)	Angry, happy, hateful, loving, sad, scared, secure	P	1/5	Singing	Swe	Tempo, timing, Int, Spectr, pitch, attack, vibrato	Classical music (lied)

Note. A dash indicates that no information was provided. Types of method included portrayal (P), synthesis (S), and manipulated portrayal (M). p = performers; l = listeners; Nat. = nationality; Spa = Spanish; Int = intensity; Art. = articulation; Swe = Swedish; Ind = Indian; Can = Canadian; Ita = Italian; Am = American; Var = various; Eng = English; Spectr = cues related to spectral energy distribution (e.g., high-frequency energy); Dut = Dutch; F0 = fundamental frequency; Jap = Japanese; Rus = Russian; Ger = German; Por = Portuguese.
[a] Acoustic cues listed within parentheses were obtained by means of listener ratings rather than acoustic measurements. [b] These terms denote particular modes of singing, which in turn are used to express different emotions.
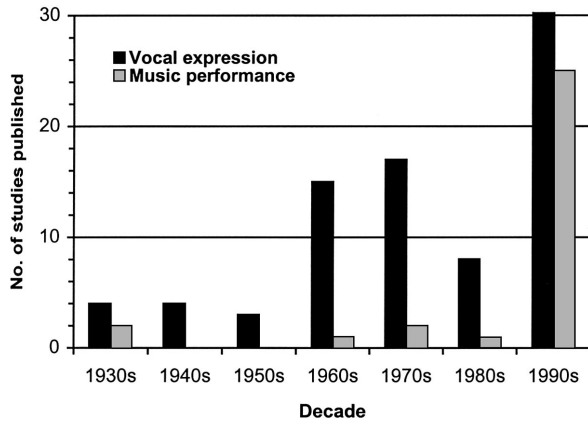
Figure 1. Number of studies of communication of emotions published for vocal expression and music performance, respectively, between 1930 and 2000.

range) of pi values, as well as confidence intervals.[6] Also indicated is the number of encoders (speakers or performers) and studies on which the estimates are based. The estimates for within-cultural vocal expression are generally based on more data than are those for cross-cultural vocal expression and music performance. As seen in Table 4, overall decoding accuracy is high for all three sets of data ($\pi = .84$–.90). Indeed, the confidence intervals suggest that decoding accuracy is typically significantly higher than what would be expected by chance alone ($\pi = .50$) for all three types of stimuli. The lowest estimate of overall accuracy in any of the 73 decoding experiments was .69 (Fenster, Blake, & Goldstein, 1977). Overall decoding accuracy across within-cultural vocal expression and music performance was .89, which is equivalent to a raw accuracy score of .70 in a forced-choice task with five response alternatives (the average number of alternatives across both channels; see, e.g., Table 1 of Rosenthal & Rubin, 1989). However, overall accuracy was significantly higher, $t(58) = 3.14$, $p < .01$, for within-cultural vocal expression ($\pi = .90$) than for cross-cultural vocal expression ($\pi = .84$). The differences in overall accuracy between music performance ($\pi = .88$) and within-cultural vocal expression and the differences between music performance and cross-cultural vocal expression were not significant. The results indicate that musical expression of emotions was about as accurate as vocal expression of emotions and that vocal expression of emotions was cross-culturally accurate, although cross-cultural accuracy was 7% lower than within-cultural accuracy in the present results. Note also that decoding accuracy for vocal expression was well above chance for both emotion portrayals and natural expressions.

The patterns of accuracy estimates for individual emotions are similar across the three sets of data. Specifically, anger ($\pi > .88$, $M = .91$) and sadness ($\pi > .91$, $M = .92$) portrayals were best decoded, followed by fear ($\pi > .82$, $M = .86$) and happiness portrayals ($\pi > .74$, $M = .82$). Worst decoded throughout was tenderness ($\pi > .71$, $M = .78$), although it must be noted that the estimates for this emotion were based on fewer data points. Further analysis confirmed that, across channels, anger and sadness were significantly better communicated (t tests, $p < .001$) than fear, happiness, and tenderness (remaining differences were not signif-

icant). This pattern of results is consistent with previous reviews of vocal expression featuring fewer studies (Johnstone & Scherer, 2000) but differs from the pattern found in studies of facial expression of emotion, in which happiness was usually better decoded than other emotions (Elfenbein & Ambady, 2002).

The standard deviation of decoding accuracy across studies was generally small, with the largest being for tenderness in music performance. (This is also indicated by the small confidence intervals for all emotions except tenderness in the case of music performance.) This finding is surprising; one would expect the accuracy to vary considerably depending on the emotions studied, the encoders, the verbal or musical material, the decoders, the procedure, and so on. Yet, the present results suggest that the estimates of decoding accuracy are fairly robust with respect to these factors. Consideration of the different measures of central tendency (unweighted mean, weighted mean, and median) shows that they differed little and that all indices gave the same patterns of findings. This suggests that the data were relatively homogenous. This impression is confirmed by plotting the distribution of data on decoding accuracy for vocal expression and music performance (see Figure 2). Only eight (11%) of the experiments yielded accuracy estimates below .80. These include three cross-cultural vocal expression experiments (two that involved natural expression), four vocal expression experiments using emotion portrayals, and one music performance experiment using drum playing as stimuli.

Possible moderators. Although the decoding data appear to be relatively homogenous, we investigated possible moderators of decoding accuracy that could explain the variability. Among the moderators were the year of the study, number of emotions encoded (this coincided with the number of response alternatives in the present data set), number of encoders, number of decoders, recording method (dummy coded, 0 = portrayal, 1 = natural sample), response format (0 = forced choice, 1 = rating scales), laboratory (dummy coded separately for Knower, Scherer, and Juslin labs; see Table 5), and channel (dummy coded separately for cross-cultural vocal expression, within-cultural vocal expression, and music performance). Table 5 presents the correlations among the investigated moderators as well as their correlations with overall decoding accuracy. Note that overall accuracy was negatively correlated with year of the study, use of natural expressions (recording method), and cross-cultural vocal expression, whereas it was positively correlated with number of emotions. The latter finding is surprising given that one would expect accuracy to decrease as the number of response alternatives increases (e.g., Rosenthal, 1982). One possible explanation is that certain earlier studies (e.g., those by Knower's laboratory) reported very high accuracy estimates (for Knower's studies, mean $\pi = .97$) although they used a large number of emotions (see Table 5). Subsequent studies featuring many emotions (e.g., Banse & Scherer, 1996) have reported lower accuracy estimates. The slightly lower overall accuracy for music performance than for within-cultural vocal expression could be related to the fact that more studies of music performance than studies of vocal expression used rating scales, which typically yield lower accuracy. In general, it is surprising

---

[6] The mean was weighted with regard to the number of encoders included.

Table 4
Summary of Results From Meta-Analysis of Decoding Accuracy for Discrete Emotions in Terms of Rosenthal and Rubin's (1989) Pi

Category	Emotion					
	Anger	Fear	Happiness	Sadness	Tenderness	Overall
Within-cultural vocal expression						
Mean (unweighted)	.93	.88	.87	.93	.82	.90
95% confidence interval	± .021	± .037	± .040	± .020	± .083	± .023
Mean (weighted)	.91	.88	.83	.93	.83	.90
Median	.95	.90	.92	.94	.85	.92
SD	.059	.095	.111	.056	.079	.072
Range	.77–1.00	.65–1.00	.51–1.00	.80–1.00	.69–.89	.69–1.00
No. of studies	32	26	30	31	6	38
No. of speakers	278	273	253	225	49	473
Cross-cultural vocal expression						
Mean (unweighted)	.91	.82	.74	.91	.71	.84
95% confidence interval	± .017	± .062	± .040	± .018		± .024
Mean (weighted)	.90	.82	.74	.91	.71	.85
Median	.90	.88	.73	.91		.84
SD	.031	.113	.077	.036		.047
Range	.86–.96	.55–.93	.61–.90	.82–.97		.74–.90
No. of studies	6	5	6	7	1	7
No. of speakers	69	66	68	71	3	71
Music performance						
Mean (unweighted)	.89	.87	.86	.93	.81	.88
95% confidence interval	± .067	± .099	± .068	± .043	± .294	± .043
Mean (weighted)	.86	.82	.85	.93	.86	.88
Median	.89	.88	.87	.95	.83	.88
SD	.094	.118	.094	.061	.185	.071
Range	.74–1.00	.69–1.00	.68–1.00	.79–1.00	.56–1.00	.75–.98
No. of studies	10	8	10	10	4	12
No. of performers	70	47	70	70	9	79

that recent studies tended to include fewer encoders, decoders, and emotions.

A simultaneous multiple regression analysis (Cohen & Cohen, 1983) with overall decoding accuracy as the dependent variable and six moderators (year of study, number of emotions, recording method, response format, Knower laboratory, and cross-cultural vocal expression) as independent variables yielded a multiple correlation of .58 (adjusted $R^2 = .27$, $F[6, 64] = 5.42$, $p < .001$; $N = 71$ with 2 outliers, standard residual $\geq 2 \Sigma$, removed). Cross-cultural vocal expression yielded a significant beta weight ($\beta = -.38$, $p < .05$), but Knower laboratory ($\beta = .20$), response format ($\beta = -.19$), recording method ($\beta = -.18$), number of emotions ($\beta = .17$), and year of the study ($\beta = .07$) did not. These results indicate that only about 30% of the variability in decoding data can be explained by the investigated moderators.[7]

Individual differences. The present results indicate that communication of emotions in vocal expression and music performance was relatively accurate. The accuracy (mean π across data sets = .87) was well beyond the frequently used criterion for correct response in psychophysical research (proportion correct [P_c] = .75), which is midway between the levels of pure guessing ($P_c = .50$) and perfect detection ($P_c = 1.00$; Gordon, 1989, p. 26). However, studies in both domains have yielded evidence of considerable individual differences in both encoding and decoding accuracy (see Banse & Scherer, 1996; Gabrielsson & Juslin, 1996; Juslin, 1997b; Juslin & Laukka, 2001; Scherer, Banse, Wallbott, & Goldbeck, 1991; Wallbott & Scherer, 1986; for a review of gender differences, see Hall, Carter, & Horgan, 2001). Particularly, en-

coders differ widely in their ability to portray specific emotions. This problem has probably contributed to the noted inconsistency of data concerning code usage in earlier research (Scherer, 1986). Because many researchers have not taken this problem seriously, several studies have investigated only one speaker or performer (see Tables 2 and 3).

Individual differences in decoding accuracy have also been reported, though they tend to be less pronounced than those in encoding. Moreover, even when decoders make incorrect responses their errors are not entirely random. Thus, error distributions are informative about the subjective similarity of various emotional expressions (Davitz, 1964a; van Bezooijen, 1984). It is of interest that the errors made in emotion decoding are similar for vocal expression and music performance. For instance, sadness and tenderness are commonly confused, whereas happiness and sadness are seldom confused (Baars & Gabrielsson, 1997; Davitz, 1964a; Davitz & Davitz, 1959; Dawes & Kramer, 1966; Fónagy, 1978; Juslin, 1997c). Similar error patterns in the two domains

---

[7] It may be argued that in many studies of vocal expression, estimates are likely to be biased because of preselection of effective portrayals before inclusion in decoding experiments. However, whether preselection of portrayals is a moderator of overall accuracy was not examined because only a minority of studies stated clearly the extent of preselection carried out. However, it should be noted that decoding accuracy of a comparable level has been found in studies that did not use preselection of emotion portrayals (Juslin & Laukka, 2001).
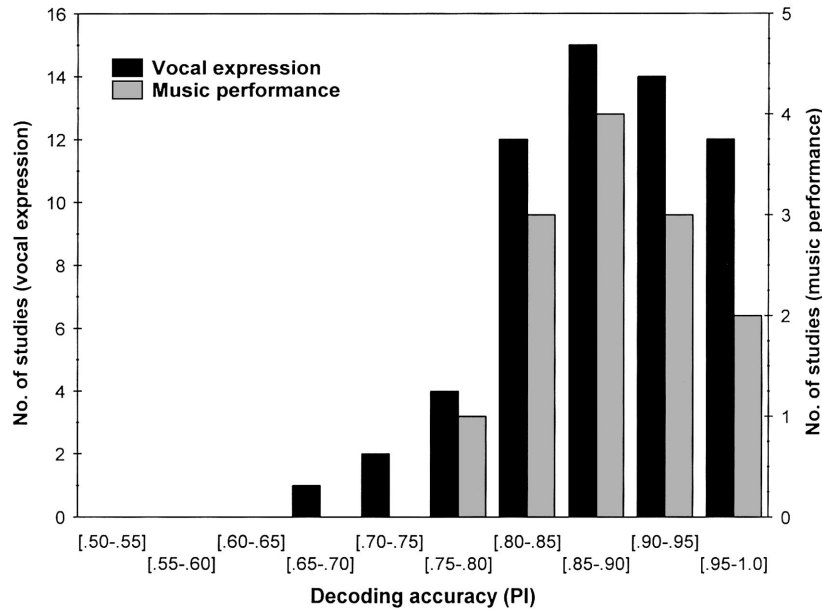
Figure 2. The distributions of point estimates of overall decoding accuracy in terms of Rosenthal and Rubin's (1989) pi for vocal expression and music performance, respectively.

provide a first indication that there could be similarities between the two channels in terms of acoustic cues.

Developmental trends. The development of the ability to decode emotions from auditory stimuli has not been well researched. Recent evidence, however, indicates that children as young as 4 years old are able to decode basic emotions from vocal expression with better than chance accuracy (Baltaxe, 1991; Friend, 2000; J. B. Morton & Trehub, 2001), at least when the verbal content is made unintelligible by using utterances in a foreign language or filtering out the verbal information (Friend, 2000).[8] The ability seems to improve with age, however, at least until school age (Dimitrovsky, 1964; Fenster et al., 1977; McCluskey & Albas, 1981; McCluskey, Albas, Niemi, Cuevas, & Ferrer, 1975) and perhaps even until early adulthood (Brosgole & Weisman, 1995; McCluskey & Albas, 1981).

Similarly, studies of music suggest that children as young as 3 or 4 years old are able to decode basic emotions from music with better than chance accuracy (Cunningham & Sterling, 1988; Dolgin & Adelson, 1990; Kastner & Crowder, 1990). Although few of these studies have distinguished between features of performance (e.g., tempo, timbre) and features of composition (e.g., mode), Dalla Bella, Peretz, Rousseau, and Gosselin (2001) found that 5-year-olds were able to use tempo (i.e., performance) but not mode (i.e., composition) to decode emotions in musical pieces. Again, decoding accuracy seems to improve with age (Adachi & Trehub, 2000; Brosgole & Weisman, 1995; Cunningham & Sterling, 1988; Terwogt & van Grinsven, 1988, 1991; but for exceptions, see Giomo, 1993; Kratus, 1993). It is interesting to note that the developmental curve over the life span appears similar for vocal expression and music but differs from that of facial expression. In a cross-sectional study, Brosgole and Weisman (1995) found that the ability to decode emotions from vocal expression and music improved during childhood and remained asymptotic

through age 43. Then, it began to decline from middle age onward (see also McCluskey & Albas, 1981). It is hard to determine whether emotion decoding occurs in children younger than 2 years old, as they are unable to talk about their experiences. However, there is preliminary evidence that infants are at least able to discriminate between some emotions in vocal and musical expressions (see Gentile, 1998; Mastropieri & Turkewitz, 1999; Nawrot, 2003; Singh, Morgan, & Best, 2002; Soken & Pick, 1999; Svejda, 1982).

Code Usage

Most early studies of vocal expression and music performance were mainly concerned with demonstrating that communication of emotions is possible at all. However, if one wants to explore communication as a process, one cannot ignore its mechanisms, in particular the code that carries the emotional meaning. A large number of studies have attempted to describe the cues used by speakers and musicians to communicate specific emotions to listeners. Most studies to date have measured only a small number of cues, but some recent studies have been more inclusive (see Tables 2 and 3). Before taking a closer look at the patterns of acoustic cues used to express discrete emotions in vocal expression and music performance, respectively, we need to consider the various cues that were used in each modality. Table 6 shows how each acoustic cue was defined and measured. The measurements were usually carried out using advanced computer software for digital analysis of speech signals. The cues extracted involve the basic dimensions of frequency, intensity, and duration, plus vari-

---

[8] This is because the verbal content may interfere with the decoding of the nonverbal content in small children (Friend, 2000).

Table 5
Intercorrelations (rs) Among Investigated Moderators of Overall Decoding Accuracy Across Vocal Expression and Music Performance

Moderators	Acc.	1	2	3	4	5	6	7	8	9	10	11	12
1. Year of the study	−.26*	—											
2. Number of emotions	.35*	−.56*	—										
3. Number of encoders	.04	−.44*	.32*	—									
4. Number of decoders	.15	−.36	.26*	−.10	—								
5. Recording method	−.24*	.14	−.24*	−.11	−.09	—							
6. Response format	−.10	.21	−.21	−.07	.09	−.07	—						
7. Knower laboratory[a]	.32*	−.67*	.50*	.50*	.36*	−.06	−.10	—					
8. Scherer laboratory[b]	−.08	.26*	−.07	−.11	.14	−.09	−.04	◊	—				
9. Juslin laboratory[c]	.02	.20	−.15	−.12	−.14	−.07	.48*	◊	◊	—			
10. Cross-cultural vocal expression	−.33*	.26*	−.11	−.18	−.08	.20	−.20	−.16	.43*	−.19	—		
11. Within-cultural vocal expression	.31*	−.41*	.34*	.22	.18	−.10	−.32*	.23*	−.22	−.28*	◊	—	
12. Music performance	−.03	.24*	−.32	−.08	−.15	−.10	.64*	−.13	−.21	.58*	◊	◊	—

Note. A diamond indicates that the correlation could not be given a meaningful interpretation because of the nature of the variables. Acc. = overall decoding accuracy.
[a] This includes the following studies: Dusenbury and Knower (1939) and Knower (1941, 1945). [b] This includes the following studies: Banse and Scherer (1996); Johnson, Emde, Scherer, and Klinnert (1986); Scherer, Banse, and Wallbott (2001); Scherer, Banse, Wallbott, and Goldbeck (1991); and Wallbott and Scherer (1986). [c] This includes the following studies: Juslin (1997a, 1997b, 1997c), Juslin and Laukka (2000, 2001), and Juslin and Madison (1999).
* p < .05.

ous combinations of these dimensions (see Table 6). For a more extensive discussion of the principles underlying production of speech and music and associated measurements, see Borden, Harris, and Raphael (1994) and Sundberg (1991), respectively. In the following, we divide data into three sets: (a) cues that are common to vocal expression and music performance, (b) cues that are specific to vocal expression, and (c) cues that are specific to music performance. The common cues are of main importance to this review, although channel-specific cues may suggest additional aspects of potential overlap that can be explored in future research.

Comparisons of common cues. Table 7 presents patterns of acoustic cues used to express different emotions as reported in 77 studies of vocal expression and 35 studies of music performance. Very few studies have reported data in such detail to permit inclusion in a meta-analysis. Furthermore, it is usually difficult to compare quantitative data across different studies because studies use different baselines (Juslin & Laukka, 2001, p. 406).[9] The most prudent approach was to summarize findings in terms of broad categories (e.g., high, medium, low), mainly according to the interpretation of the authors of each study but (whenever possible) with support from actual data provided in tables and figures. Many studies provided only partial reports of data or reported data in a manner that required careful analysis to extract usable data points. In the few cases in which we were uncertain about the interpretation of a particular data point, we simply omitted this data point from the review. In the majority of cases, however, the scoring of data was straightforward, and we were able to include 1,095 data points in the comparisons.

Starting with the cues used freely in both channels (i.e., in vocal expression/music performance, respectively: speech rate/tempo, voice intensity/sound level, and high-frequency energy), there are relatively similar patterns of cues for the two channels (see Table 7). For example, speech rate/tempo and voice intensity/sound level were typically increased in anger and happiness, whereas they were decreased in sadness and tenderness. Furthermore, the high-frequency energy was typically increased in happiness and anger,

whereas it was decreased in sadness and tenderness. Although less data were available concerning voice intensity/sound level variability, the results were largely similar for the two channels. The variability increased in anger and fear but decreased in sadness and tenderness. However, there were differences as well. Note that fear was most commonly associated with high intensity in vocal expression, albeit low intensity (sound level) in music performance. One possible explanation of this inconsistency is that the results reflect various intensities of the same emotion (e.g., strong fear and weak fear) or qualitative differences among closely related emotions. For example, mild fear may be associated with low voice intensity and little high-frequency energy, whereas panic fear may be associated with high voice intensity and much high-frequency energy (Banse & Scherer, 1996; Juslin & Laukka, 2001). Thus, it is possible that studies of music performance have studied almost exclusively mild fear, whereas studies of vocal expression have studied both mild fear and panic fear. This explanation is clearly consistent with the present results when one considers our findings of bimodal distributions of intensity and high-frequency energy for vocal expression as compared with unimodal distributions of sound level and high-frequency energy for music performance (see Table 7). However, to confirm this interpretation one would need to conduct studies of music performance that systematically manipulate emotion intensity in expressions of fear. Such studies are currently underway (e.g., Juslin & Lindström, 2003). Overall, however, the results were relatively similar across the channels for the three major cues (speech rate/tempo, vocal intensity/sound level, and high-frequency energy), although there were relatively

---

[9] Baseline refers to the use of some kind of frame of reference (e.g., a neutral expression or the average across emotions) against which emotion-specific changes in acoustic cues are indexed. The problem is that many types of baseline (e.g., the average) are sensitive to what emotions were included in the study, which renders studies that included different emotions incomparable.

Table 6
Definition and Measurement of Acoustic Cues in Vocal Expression and Music Performance

Acoustic cues	Perceived correlate	Definition and measurement
		Vocal expression
Pitch		
Fundamental frequency (F0)	Pitch	F0 represents the rate at which the vocal folds open and close across the glottis. Acoustically, F0 is defined as the lowest periodic cycle component of the acoustic waveform, and it is extracted by computerized tracking algorithms (Scherer, 1982).
F0 contour	Intonation contour	The F0 contour is the sequence of F0 values across an utterance. Besides changes in pitch, the F0 contour also contains temporal information. The F0 contour is hard to operationalize, and most studies report only qualitative classifications (Cowie et al., 2001).
Jitter	Pitch perturbations	Jitter is small-scale perturbations in F0 related to rapid and random fluctuations of the time of the opening and closing of the vocal folds from one vocal cycle to the next. Extracted by computerized tracking algorithms (Scherer, 1989).
Intensity		
Intensity	Loudness of speech	Intensity is a measure of energy in the acoustic signal, and it reflects the effort required to produce the speech. Usually measured from the amplitude acoustic waveform. The standard unit used to quantify intensity is a logarithmic transform of the amplitude called the *decibel* (dB; Scherer, 1982).
Attack	Rapidity of voice onsets	The attack refers to the rise time or rate of rise of amplitude for voiced speech segments. It is usually measured from the amplitude acoustic waveform (Scherer, 1989).
Temporal aspects		
Speech rate	Velocity of speech	The rate can be measured as overall duration or as units per duration (e.g., words per min). It may include either complete utterances or only the voiced segments of speech (Scherer, 1982).
Pauses	Amount of silence in speech	Pauses are usually measured as number or duration of silences in the acoustic waveform (Scherer, 1982).
Voice quality		
High-frequency energy	Voice quality	High-frequency energy refers to the relative proportion of total acoustic energy above versus below a certain cut-off frequency (e.g., Scherer et al., 1991). As the amount of high-frequency energy in the spectrum increases, the voice sounds more sharp and less soft (Von Bismarck, 1974). It is obtained by measuring the long-term average spectrum, which is the distribution of energy over a range of frequencies, averaged over an extended time period.
Formant frequencies	Voice quality	Formant frequencies are frequency regions in which the amplitude of acoustic energy in the speech signal is high, reflecting natural resonances in the vocal tract. The first two formants largely determine vowel quality, whereas the higher formants may be speaker dependent (Laver, 1980). The mean frequency and the width of the spectral band containing significant formant energy are extracted from the acoustic waveform by computerized tracking algorithms (Scherer, 1989).
Precision of articulation	Articulatory effort	The vowel quality tends to move toward the formant structure of the neutral schwa vowel (e.g., as in *sofa*) under strong emotional arousal (Tolkmitt & Scherer, 1986). The precision of articulation can be measured as the deviation of the formant frequencies from the neutral formant frequencies.
Glottal waveform	Voice quality	The glottal flow waveform represents the time air is flowing between the vocal folds (abduction and adduction) and the time the glottis is closed for each vibrational cycle. The shape of the waveform helps to determine the loudness of the sound generated and its timbre. A jagged waveform represents sudden changes in airflow that produce more high frequencies than a soft waveform. The glottal waveform can be inferred from the acoustical signal using inverse filtering (Laukkanen et al., 1996).
		Music performance
Pitch		
F0	Pitch	Acoustically, F0 is defined as the lowest periodic cycle component of the acoustic waveform. One can distinguish between the macro pitch level of particular musical pieces, and the micro intonation of the performance. The former is often given in the unit of the semitone, the latter is given in terms of deviations from the notated macro pitch (e.g., in cents; Sundberg, 1991).
F0 contour	Intonation contour	F0 contour is the sequence of F0 values. In music, intonation refers to manner in which the performer approaches and/or maintains the prescribed pitch of notes in terms of deviations from precise pitch (Baroni et al., 1997).

Table 6 (*continued*)

Acoustic cues	Perceived correlate	Definition and measurement
		Music performance (*continued*)
Vibrato	Vibrato	Vibrato refers to periodic changes in the pitch (or loudness) of a tone. Depth and rate of vibrato can be measured manually from the F0 trace (or amplitude envelope; Metfessel, 1932).
Intensity		
Intensity	Loudness	Intensity is a measure of the energy in the acoustic signal. It is usually measured from the amplitude of the acoustic waveform. The standard unit used to quantify intensity is a logarithmic transformation of the amplitude called the *decibel* (dB; Sundberg, 1991).
Attack	Rapidity of tone onsets	Attack refers to the rise time or rate of rise of the amplitude of individual notes. It is usually measured from the acoustic waveform (Kotlyar & Morozov, 1976).
Temporal aspects		
Tempo	Velocity of music	The mean tempo of a performance is obtained by dividing the total duration of the performance until the onset of its final note by the number of beats and then calculating the number of beats per min (bpm; Bengtsson & Gabrielsson, 1980).
Articulation[a]	Proportion of sound to silence in successive notes	The mean articulation of a performance is typically obtained by measuring two durations for each tone—the duration from the onset of a tone until the onset of the next tone (d_{ii}), and the duration from the onset of a tone until its offset (d_{io}). These durations are used to calculate the $d_{io}{:}d_{ii}$ ratio (the articulation) of each tone (Bengtsson & Gabrielsson, 1980). These values are averaged across the performance and expressed as a percentage. A value around 100% refers to *legato* articulation; a value of 70% or lower refers to *staccato* articulation (Woody, 1997).
Timing	Tempo and rhythm variation	Timing variations are usually described as *deviations* from the nominal values of a musical notation. Overall measures of the amount of deviations in a performance may be obtained by calculating the number of notes whose deviation is less than a given percentage of the note value. Another index of timing changes concerns so-called durational contrasts between long and short notes in rhythm patterns. Contrasts may be played with "sharp" durational contrasts (close to or larger than the nominal ratio) or with "soft" durational contrasts (a reduced ratio; Gabrielsson, 1995).
Timbre		
High-frequency energy	Timbre	High-frequency energy refers to the relative proportion of total acoustic energy above versus below a certain cut-off frequency in the frequency spectrum of the performance (Juslin, 2000). In music, timbre is in part a characteristic of the specific instrument. However, different techniques of playing may also influence the timbre of many instruments, such as the guitar (Gabrielsson & Juslin, 1996).
Singer's formant	Timbre	The singer's formant refers to a strong resonance around 2500–3000 Hz and is that which adds brilliance and carrying power to the voice. It is attributed to a lowered larynx and widened pharynx, which forms an additional resonance cavity (Sundberg, 1999).

[a] This use of the term *articulation* should be distinguished from its use in studies of vocal expression, where *articulation* refers to the settings of the articulators (lips, tongue, lower jaw, pharyngeal sidewalls) that determine the resonant characteristics of the vocal tract (Sundberg, 1999). To avoid confusion in this review, we use the term *articulation* only in its musical sense, whereas vocal expression articulation is considered only in terms of its consequences for voice quality (e.g., in the term *precision of articulation*).

few data points in some of the comparisons. It must be noted that these cues account for a large proportion of the variance in listeners' judgments of emotional expression in synthesized sound sequences (e.g., Juslin, 1997c; Juslin & Madison, 1999; Scherer & Oshinsky, 1977). The present results are not as clear cut as one would hope for, but some inconsistency in the data is only what should be expected given that there were large individual differences among encoders and that many studies included only a single encoder. (Further explanation of the inconsistency in these findings is provided in the Discussion section.)

In addition to the converging findings for the three major cues, there are also similarities with regard to some other cues. It can be seen in Table 7 that the degree to which emotion portrayals display microstructural regularity versus irregularity (with respect to frequency, intensity, and duration) can discriminate between certain emotions. Specifically, it would appear that positive emotions (happiness, tenderness) are more regular than negative emotions (anger, fear, sadness). That is, irregularities in frequency, intensity, and duration seem to be signs of negative emotion. This hypothesis, mentioned by Davitz (1964a), deserves more attention in future research.

Another cue that has been little studied so far is voice onsets/tone attack. Sundberg (1999) observed that perceptual stimuli that change are easier to process than quasi-stationary stimuli and that the beginning and the end of a sound may be particularly revealing. Indeed, the limited data available suggest that voice onsets and tone attacks differed depending on the emotion expressed. As can be seen in Table 7, studies of music performance suggest that fast tone attacks were used in anger and happiness, whereas slow tone attacks were used in sadness and tenderness.

Table 7
Patterns of Acoustic Cues Used to Express Discrete Emotions in Studies of Vocal Expression and Music Performance

Emotion	Category	Vocal expression studies	Category	Music performance studies
		Speech rate		Tempo
Anger	**Fast**	(1, 4, 6, 9, 13, 16, 18, 19, 24, 27, 28, 42, 48, 55, 64, 69, 70, 71, 72, 80, 82, 83, 89, 92, 97, 99, 100, 103)	**Fast**	(2, 4, 7, 8, 9, 10, 13, 14, 16, 18, 19, 20, 22, 23, 25, 26, 29, 34, 37, 41)
	Medium	(51, 63, 75)	Medium	(27)
	Slow	(6, 10, 47, 53)	Slow	
Fear	**Fast**	(3, 4, 9, 10, 13, 16, 18, 28, 33, 47, 48, 51, 63, 64, 71, 72, 80, 82, 83, 89, 95, 96, 97, 103)	**Fast**	(2, 8, 10, 23, 25, 26, 27, 32, 37)
	Medium	(1, 53, 70)	Medium	(18, 20, 22)
	Slow	(12, 42)	Slow	(7, 19)
Happiness	**Fast**	(4, 6, 9, 10, 16, 19, 20, 24, 33, 37, 42, 53, 55, 71, 72, 75, 80, 82, 83, 97, 99, 100)	**Fast**	(1, 2, 3, 4, 7, 8, 9, 12, 13, 14, 15, 16, 18, 19, 20, 22, 23, 26, 27, 29, 34, 37, 41)
	Medium	(13, 18, 51, 89, 92)	Medium	(25, 32)
	Slow	(1, 3, 48, 64, 70, 95)	Slow	
Sadness	Fast	(95)	Fast	
	Medium	(1, 51, 53, 63, 70)	Medium	
	Slow	(3, 4, 6, 9, 10, 13, 16, 18, 19, 20, 24, 27, 28, 37, 48, 55, 64, 69, 71, 72, 75, 80, 82, 83, 89, 92, 97, 99, 100, 103)	**Slow**	(2, 3, 4, 7, 8, 9, 10, 13, 15, 16, 18, 19, 20, 21, 22, 23, 25, 26, 27, 29, 32, 34, 37)
Tenderness	Fast	(4)	Fast	
	Medium		Medium	
	Slow	(24, 33, 96)	**Slow**	(1, 2, 7, 8, 10, 13, 14, 16, 19, 27, 41)

Emotion	Category	Vocal expression studies	Category	Music performance studies
		Voice intensity (*M*)		Sound level (*M*)
Anger	**High**	(1, 3, 6, 8, 9, 10, 18, 24, 27, 33, 42, 43, 44, 47, 50, 51, 55, 56, 61, 63, 64, 70, 72, 82, 89, 91, 95, 97, 99, 101)	**High**	(2, 5, 7, 8, 9, 13, 14, 15, 16, 18, 19, 20, 22, 23, 25, 26, 27, 29, 32, 34, 35, 36, 41)
	Medium	(4)	Medium	
	Low	(79)	Low	
Fear	**High**	(3, 4, 6, 18, 45, 47, 50, 63, 79, 82, 89)	High	
	Medium	(70, 95, 97)	Medium	(27)
	Low		Low	
Happiness	Low	(9, 10, 33, 42, 43, 51, 56, 64)	Low	(2, 7, 13, 18, 19, 20, 22, 23, 25, 26, 32)
	High	(3, 4, 6, 8, 10, 24, 43, 44, 45, 50, 51, 52, 56, 72, 82, 88, 89, 97, 99, 101)	High	(1, 2, 7, 13, 14, 26, 41)
	Medium	(18, 42, 55, 70, 91, 95)	**Medium**	(5, 8, 16, 18, 19, 20, 22, 23, 25, 27, 29, 32, 34)
	Low		Low	(9)
Sadness	High	(79)	High	
	Medium	(4, 95)	Medium	(18, 25)
	Low	(1, 3, 6, 8, 9, 10, 18, 24, 27, 44, 45, 47, 50, 51, 52, 55, 56, 61, 63, 64, 70, 72, 82, 88, 89, 91, 97, 99, 101)	**Low**	(5, 7, 8, 9, 13, 15, 16, 19, 20, 21, 22, 23, 26, 27, 28, 29, 32, 34, 36)
Tenderness	High		High	
	Medium		Medium	
	Low	(4, 24, 33, 95)	**Low**	(1, 7, 8, 12, 13, 14, 16, 19, 27, 28, 41)

Emotion	Category	Vocal expression studies	Category	Music performance studies
		Voice intensity variability		Sound level variability
Anger	**High**	(51, 61, 64, 70, 80, 82, 89, 95, 101)	**High**	(14, 29, 41)
	Medium	(3)	Medium	
	Low	(10, 53)	Low	(23, 34)

Table 7 (continued)

Emotion	Category	Vocal expression studies (continued)	Category	Music performance studies (continued)
		Voice intensity variability (continued)		**Sound level variability (continued)**
Fear	**High**	(10, 64, 68, 80, 82, 83, 95)	**High**	(8, 10, 13, 23, 37)
	Medium	(3, 51, 53, 70)	Medium	
	Low	(89)	Low	
Happiness	**High**	(3, 10, 53, 64, 80, 82, 95, 101)	High	(1, 14, 34)
	Medium	(51, 70, 89)	Medium	(41)
	Low	(47, 83)	**Low**	(8, 23, 29, 37)
Sadness	High	(10, 89)	High	(23)
	Medium		Medium	
	Low	(53)	**Low**	(29, 34)
Tenderness	High	(3, 51, 61, 64, 70, 82, 95, 101)	High	
	Medium		Medium	
	Low		**Low**	(1, 14, 41)
		High-frequency energy		**High-frequency energy**
Anger	**High**	(4, 6, 16, 18, 24, 33, 35, 38, 42, 47, 50, 51, 54, 56, 63, 65, 73, 82, 83, 91, 97, 103)	**High**	(8, 10, 13, 15, 16, 19, 20, 22, 26, 36, 37)
	Medium		Medium	
	Low		Low	
Fear	**High**	(18, 33, 54, 56, 63, 83, 97, 103)	High	(10, 37)
	Medium	(4, 6)	Medium	
	Low	(16, 38, 42, 50, 51, 82)	**Low**	(15, 19, 20, 22, 26, 36)
Happiness	**High**	(4, 42, 50, 51, 52, 54, 56, 73, 82, 83, 88, 91, 97)	High	(15)
	Medium	(6, 18, 24)	**Medium**	(8, 10, 12, 13, 16, 19, 20, 22, 26, 36)
	Low	(83)	Low	(37)
Sadness	High		High	
	Medium		Medium	
	Low	(4, 6, 16, 18, 24, 38, 50, 51, 52, 54, 56, 63, 73, 82, 83, 88, 91, 97, 103)	**Low**	(8, 10, 19, 20, 22, 26, 35, 36, 37)
Tenderness	High		High	
	Medium		Medium	
	Low	(4, 24, 33)	**Low**	(8, 10, 12, 13, 16, 19)
		F0 (M)		**F0 (M)**[a]
Anger	**High**	(3, 6, 10, 14, 16, 19, 24, 27, 29, 32, 34, 36, 45, 46, 48, 51, 55, 61, 63, 64, 65, 71, 72, 73, 74, 79, 82, 83, 95, 97, 99, 101, 103)	**Sharp**	(26, 37)
	Medium	(4, 33, 50, 70, 75)	Precise	(32)
	Low	(18, 43, 44, 53, 89)	Flat	(8)
Fear	**High**	(1, 4, 6, 12, 16, 17, 18, 29, 43, 45, 47, 48, 50, 53, 60, 63, 65, 70, 71, 72, 74, 82, 83, 87, 89, 93, 97, 102)	**Sharp**	(32, 37)
	Medium	(3, 10, 32, 47, 51, 95, 96, 103)	Precise	
	Low	(3, 64, 79)	Flat	(8)
Happiness	**High**	(1, 3, 4, 6, 10, 14, 16, 19, 20, 24, 32, 33, 37, 43, 44, 45, 46, 47, 50, 51, 52, 53, 55, 64, 71, 74, 75, 80, 82, 86, 87, 88, 97, 99)	**Sharp**	(8, 26, 32)
			High	(Hevner, 1937; Rigg, 1940; Wedin, 1972)
	Medium	(70, 101)	Precise	
	Low	(18, 89)	Flat	
Sadness	High	(16, 53, 70, 97)	Sharp	
	Medium	(18)	Precise	

(table continues)

Table 7 (*continued*)

Emotion	Category	Vocal expression studies	Category	Music performance studies
		F0 (*M*) (*continued*)		**F0 (*M*)[a] (*continued*)**
	Low	(3, 4, 6, 10, 14, 16, 19, 20, 24, 27, 29, 32, 37, 44, 45, 46, 47, 50, 51, 52, 55, 61, 63, 64, 71, 72, 73, 74, 75, 79, 82, 83, 86, 88, 89, 92, 95, 99, 101, 103)	**Flat**	(4, 13, 16, 26, 32, 37)
			Low	(Gundlach, 1935; Hevner, 1937; Rigg, 1940; K. B. Watson, 1942; Wedin, 1972)
Tenderness	High	(33)	Sharp	
	Medium		Precise	
	Low	(4, 24, 32, 96)	Flat	
		F0 variability		**F0 variability**
Anger	**High**	(1, 4, 6, 9, 13, 14, 16, 18, 24, 29, 33, 36, 42, 47, 51, 55, 56, 63, 70, 71, 72, 74, 89, 95, 97, 99, 103)	**High**	(32)
	Medium	(3, 64, 75, 82)	Medium	(3)
	Low	(10, 44, 50, 83)	Low	(37)
Fear	**High**	(4, 13, 16, 18, 29, 56, 80, 89, 103)	High	(32)
	Medium	(47, 72, 74, 82, 95, 97)	Medium	
Happiness	**Low**	(1, 3, 6, 10, 32, 33, 42, 47, 50, 51, 63, 64, 70, 71, 80, 83, 96)	**Low**	(12, 37)
	High	(3, 4, 6, 9, 10, 13, 16, 18, 20, 32, 33, 37, 42, 44, 47, 48, 50, 51, 55, 56, 64, 71, 72, 74, 75, 80, 82, 83, 86, 89, 95, 97, 99)	**High**	(3, 12, 37)
	Medium	(14, 70)	Medium	
Sadness	Low	(1)	Low	(32)
	High	(10, 20)	High	
	Medium	(6)	Medium	
	Low	(3, 4, 9, 13, 14, 16, 18, 29, 32, 37, 44, 47, 50, 51, 55, 56, 63, 64, 70, 71, 72, 74, 75, 80, 82, 86, 89, 95, 97, 99, 103)	**Low**	(3, 29, 32)
Tenderness	High		High	
	Medium		Medium	
	Low	(4, 24, 33, 95, 96)	**Low**	(12)
		F0 contours		**Pitch contours**
Anger	**Up**	(30, 32, 33, 51, 83, 103)	**Up**	(12, 37)
	Down	(9, 36)	Down	
Fear	**Up**	(9, 18, 51, 74, 80, 83)	**Up**	(12, 37)
	Down		Down	
Happiness	**Up**	(18, 20, 24, 32, 33, 51, 52)	**Up**	(1, 8, 12, 35, 37)
	Down		Down	
Sadness	Up		Up	
	Down	(20, 24, 30, 51, 52, 63, 72, 74, 83, 86, 103)	**Down**	(8, 37)
Tenderness	Up	(24)	Up	
	Down	(32, 33, 96)	**Down**	(12)
		Voice onsets		**Tone attack**
Anger	Fast	(83)	**Fast**	(4, 13, 16, 18, 19, 25, 29, 35, 41)
	Slow	(51)	Slow	
Fear	Fast	(51)	Fast	(25)
	Slow	(83)	**Slow**	(19, 37)

Table 7 (*continued*)

Emotion	Category	Vocal expression studies — Voice onsets (*continued*)	Category	Music performance studies — Tone attack (*continued*)
Happiness	**Fast**	(51, 83)	**Fast**	(13, 19, 25, 29, 37)
	Slow		Slow	(4)
Sadness	Fast	(51)	Fast	
	Slow	(83)	**Slow**	(4, 10, 13, 16, 19, 25, 29, 37)
Tenderness	Fast		Fast	
	Slow	(31)	**Slow**	(10, 13, 19)

Emotion	Category	Vocal expression studies — Microstructural regularity	Category	Music performance studies — Microstructural regularity
Anger	Regular		Regular	(4)
	Irregular	(8, 24, 103)	**Irregular**	(8, 13, 14, 28)
Fear	Regular		Regular	(28)
	Irregular	(30, 103)	**Irregular**	(8, 10, 13)
Happiness	**Regular**	(8, 24)	**Regular**	(4, 8, 14, 28)
	Irregular		Irregular	
Sadness	Regular		Regular	(28)
	Irregular	(8, 24, 27, 103)	**Irregular**	(5, 8)
Tenderness	**Regular**	(24)	**Regular**	(8, 14, 28)
	Irregular		Irregular	

Note. Numbers within parantheses refer to studies, as indicated in Table 2 (vocal expression) and Table 3 (music performance) respectively. Text in bold indicates the most frequent finding for respective acoustic cue and modality. F0 = fundamental frequency.

[a] As regards F0 (*M*) in music, one should distinguish between the micro intonation of the performance, which may be sharp (higher than prescribed pitch), precise (prescribed pitch), or flat (below prescribed pitch) and the macro pitch level (e.g., high, low) of specific pieces of music (see Table 6). The additional studies cited here focus on the latter aspect.

The data for studies of vocal expression are less convincing. However, only three studies have reported data on voice onsets thus far, and these studies used different methods (synthesized sound sequences and listener judgments in Scherer & Oshinsky, 1977, vs. measurements of emotion portrayals in Fenster et al., 1977, and Juslin & Laukka, 2001).

In this review, we chose to concentrate on those features that may be independently controlled by speakers and performers almost regardless of the verbal and musical material used. As we noted, a number of variables in musical compositions (e.g., harmony, scales, mode) do not have any direct counterpart in vocal expression, and vice versa. However, if we broaden the perspective for one moment, there is actually one aspect of musical compositions that has an approximate counterpart in vocal expression, namely, the pitch level. The pitch level in musical compositions might be compared with the fundamental frequency (F0) in vocal expression. Thus, it is interesting to note that low pitch was associated with sadness in both vocal expression and musical compositions (see Table 7), whereas high pitch was associated with happiness in both vocal expression and musical compositions (for a review of the latter, see Gabrielsson & Juslin, 2003). The cross-modal evidence for the other three emotions is still provisionary, although studies of vocal expression suggest that anger and fear are primarily associated with high pitch, whereas tenderness is primarily associated with low pitch. A study by Patel, Peretz, Tramo, and Labreque (1998) suggests that the same neural resources may be involved in the processing of F0 contours in speech and melody contours in music; thus, there may also be similarities regarding pitch contours. For example, rising F0 contours may be associated with "active" emotions (e.g., happiness, anger, fear), whereas falling contours may be associated with less active emotions (e.g., sadness, tenderness; Cordes, 2000; Fónagy, 1978; M. Papoušek, 1996; Scherer & Oshinsky, 1977; Sedláček & Sychra, 1963). This hypothesis is supported by the present data (Table 7) in that anger, fear, and happiness were associated with a higher proportion of upward pitch contours than were sadness and tenderness. Further research is needed to confirm these preliminary results, because most of the data were based on informal observations or simple acoustic indices that do not capture the complex nature of F0 contours in vocal expression. (For an attempt to develop a more sensitive measure of F0 contour using curve fitting, see Katz, Cohn, & Moore, 1996.)

Cues specific to vocal expression.   Table 8 presents additional data for acoustic cues measured specifically in vocal expression. These cues, by and large, have not been investigated systematically, but some tendencies can still be observed. For instance, there is fairly strong evidence that portrayals of sadness involved a large proportion of pauses, whereas portrayals of anger involved a small proportion of pauses. (Note that the former relationship has been considered as an acoustic correlate to depression; see Ellgring & Scherer, 1996.) The data were less consistent for portrayals of fear and happiness, and pause distributions in tenderness have been little studied thus far. As regards measurements of formant frequencies, the results were, again, most consistent for anger and sadness; beginning with precision of articulation, Table 8 shows that anger was associated with increases in precision of articulation, whereas sadness was associated with decreases. Similarly, results so far indicate that Formant 1 (F1) was raised in anger and happiness but lowered in fear and sadness. Furthermore, the data

indicate that F1 bandwidth (bw) was narrowed in anger and happiness but widened in fear and sadness. Clearly, though, these findings must be regarded as preliminary.

It should be noted that the results for F1, F1 (bw), and precision of articulation may be partly explained by intercorrelations that reflect the underlying vocal production (Borden et al., 1994). A tense voice leads to pharyngeal constriction and tensing, as well as a shortening of the vocal tract, which leads to a rise in F1, a more narrow F1 (bw), and stronger high-frequency resonances. This pattern was seen in anger portrayals. Sadness portrayals, on the other hand, appear to have involved a lax voice, with unconstricted pharynx and lower subglottal pressure, which yields lower F1, precision of articulation, and high-frequency but wider F1 (bw). It has been found that formant frequencies can be affected by facial expression. Smiling tends to raise formant frequencies (Tartter, 1980), whereas frowning tends to lower them (Tartter & Braun, 1994). A few studies have measured glottal waveform (see, in particular, Laukkanen, Vilkman, Alku, & Oksanen, 1996), and again the results were most consistent for portrayals of anger and sadness: Anger was associated with steep glottal waveforms, whereas sadness was associated with rounded waveforms. Results regarding jitter (F0 perturbation) are still preliminary, partly because of the problems involved in reliably measuring jitter. As seen in Table 8, the results do not yet allow any definitive conclusions other than the possible tendency for anger portrayals to show more jitter than sadness portrayals. It is quite possible that jitter is a voice cue that is difficult to manipulate for actors and therefore that more consistent results require the use of natural samples of vocal expression (Bachorowski & Owren, 1995).

Cues specific to music performance.   Table 9 shows additional data for acoustic cues measured specifically in music performance. One of the fundamental cues in music performance is articulation (i.e., the relative proportion of sound to silence in note values; see Table 6). *Staccato articulation* means that there is much air between the notes, whereas *legato articulation* means that the notes are played continuously. The results concerning articulation were relatively consistent. Anger, fear, and happiness were associated primarily with staccato articulation, whereas sadness and tenderness were associated primarily with legato articulation. One exception is that guitar players tended to play anger with legato articulation (see, e.g., Juslin, 1993, 1997b, 2000), suggesting that the code is not entirely invariant across musical instruments. Both the mean value and the standard deviation of the articulation can be important, although the two are intercorrelated to some extent (Juslin, 2000) such that when the articulation becomes more staccato the variability increases as well. This is explained by the fact that certain notes in the musical structure are performed legato regardless of the expression. Therefore, when the remaining notes are played staccato, the variability automatically increases. However, this intercorrelation is not perfect. For instance, anger and happiness expressions were both associated with staccato mean articulation, but only happiness expressions were associated with large articulation variability (Juslin & Madison, 1999; see Table 9). Closer study of the patterns of articulation within musical performances may provide important clues about characteristics associated with various emotions (Juslin & Madison, 1999; Madison, 2000b).

The data regarding use of vibrato (i.e., periodic changes in the

pitch of a tone) were relatively inconsistent (see Table 9) and suggest that music performers did not use vibrato systematically to communicate particular emotions. Large vibrato extent in anger portrayals and slow vibrato rate in sadness portrayals were the only consistent tendencies, with the possible addition of fast vibrato rate in fear and happiness. It is still possible that extent and rate of vibrato is consistently related to listeners' judgments of emotion because it has been shown that listeners can correctly decode emotions like anger, fear, happiness, and sadness from single notes that feature vibrato (Konishi, Imaizumi, & Niimi, 2000).

Because music is usually performed according to a metrical framework, it is meaningful to describe the nature of a performance in terms of its microstructural deviations from prescribed note values (Gabrielsson, 1999). Data concerning timing variability suggest that fear portrayals showed most timing variability, followed by anger, sadness, and tenderness portrayals. Happiness portrayals showed the least timing variability of all. Moreover, limited findings regarding durational contrasts between long and short notes indicate that the contrasts were increased (sharp) in anger and fear portrayals, whereas they were reduced (soft) in sadness and tenderness portrayals. The results for happiness portrayals were still equivocal. Finally, a few studies measured the singer's formant as a function of emotional expression, although more data are needed before any definitive conclusions can be drawn (see Table 9).

Relative importance of different cues. What is the relative importance of the different acoustic cues in vocal expression and music performance? The findings from a number of studies have shown that speech rate/tempo, voice intensity/sound level, voice quality/timbre, and F0/pitch are among the most powerful cues in terms of their effects on listeners' ratings of emotional expression (Juslin, 1997c, 2000; Juslin & Madison, 1999; Lieberman & Michaels, 1962; Scherer & Oshinsky, 1977). In particular, studies that used synthesized sound sequences indicate that speech rate/tempo was of primary importance for listeners' judgments of emotional expression (Juslin, 1997c; Scherer & Oshinsky, 1977; see also Gabrielsson & Juslin, 2003), but in music performance, the impact of tempo was decreased if listeners were required to judge different melodies with different associated baselines of tempo (e.g., Juslin, 2000). Similar effects may also occur with regard to different baselines of speech rate for different speakers.

It is interesting to note that when researchers of nonverbal communication of emotion have investigated how people use various nonverbal channels to infer emotional states in everyday life, they most frequently report vocal cues. They particularly mention using loudness and speed of talking (e.g., Planalp, 1998), the same cues (i.e., sound level and tempo) that explain most variance in listeners' judgments of emotional expression in musical performances (Juslin, 1997c, 2000; Juslin & Madison, 1999). There is further indication that cue levels (e.g., mean tempo) have a larger influence on listeners' judgments than do patterns of cue variability (e.g., timing patterns; Figure 1 of Madison, 2000a).

Comparison with Scherer's (1986) predictions. Table 10 presents a comparison of the summarized findings in vocal expression and music performance with Scherer's (1986) theoretical predictions. Because of the problems associated with establishing a precise baseline, we compare results and predictions simply in

terms of direction of effect rather than in terms of specific degrees of effect. Table 10 shows the data for eight voice cues and four emotions (Scherer, 1986, did not make predictions for love–tenderness), for a total of 32 comparisons. The comparisons are made in regard to Scherer's (1986) predictions for rage–hot anger, fear–terror, elation–joy, and sadness–dejection because these correspond best, in our view, with the emotions most frequently investigated. Careful inspection of Table 10 reveals that 27 (84%) of the predictions match the present results. Predictions and results did not match in the cases of F0 (*SD*) and F1 (*M*) for fear, as well as in the case of F1 (*M*) for happiness. However, the findings are generally consistent with Scherer's (1986) physiologically based predictions.

Discussion

The empirical findings reviewed in this article generally support the theoretical predictions made at the outset. First, it is clear that communication of emotions may reach an accuracy well above the accuracy that would be expected by chance alone in both vocal expression and music performance—at least for broad emotion categories corresponding to basic emotions (i.e., anger, sadness, happiness, fear, love). Decoding accuracy for individual emotions showed similar patterns for the two channels. Anger and sadness were generally better communicated than fear, happiness, and tenderness. Second, the findings indicate that vocal expression of emotion was cross-culturally accurate, although the accuracy was lower than for within-cultural vocal expression. Unfortunately, relevant data with regard to music performance are still lacking. Third, there is preliminary evidence that the ability to decode basic emotions from vocal expression and music performance develops in early childhood at least, perhaps even in infancy. Fourth, the present findings strongly suggest that music performance uses largely the same emotion-specific patterns of acoustic cues as does vocal expression. Table 11 presents the hypothesized emotion-specific patterns of cues according to this review, which could be subjected to direct tests in listening experiments using synthesized and systematically varied sound sequences.[10] However, the review has also revealed many gaps in the data base that must be filled in further research (see Tables 7–9). Finally, the emotion-specific patterns of acoustic cues were mainly consistent with Scherer's (1986) predictions, which presumed a correspondence between emotion-specific physiological changes and voice production.[11] Taken together, these findings, which are

---

[10] It may be noted that the pattern of cues for sadness is fairly similar to the pattern of cues obtained in studies of vocal correlates of clinical depression (see Alpert, Pouget, & Silva, 2001; Ellgring & Scherer, 1996; Hargreaves et al., 1965; Kuny & Stassen, 1993; Nilsonne, 1987; Stassen, Kuny, & Hell, 1998).

[11] Note that these findings are consistent both with basic emotions theory and component process theory in showing that there are emotion-specific patterning of acoustic cues over and above what would be predicted by a dimensional approach involving the dimensions activation and valence. However, these studies did not test the most important of the component theory's assumptions, namely that there are highly differentiated, sequential patterns of cues that reflect the cumulative result of the adaptive changes produced by a specific appraisal profile (Scherer, 2001). See Johnstone (2001) for an attempt to test this notion.

Table 8
Patterns of Acoustic Cues Used to Express Emotions Specifically in Vocal Expression Studies

Emotion	Category	Vocal expression studies
	Proportion of pauses	
Anger	Large	
	Medium	
	Small	(4, 18, 27, 28, 47, 51, 89, 95)
Fear	Large	(4, 27)
	Medium	(18, 51, 95)
	Small	(28, 30, 47, 89)
Happiness	Large	(64)
	Medium	(51, 89)
	Small	(4, 18, 47)
Sadness	**Large**	(1, 4, 18, 24, 28, 47, 51, 72, 89, 95, 103)
	Medium	
	Small	(27)
Tenderness	**Large**	(4)
	Medium	
	Small	
	Precision of articulation	
Anger	**High**	(16, 18, 24, 51, 54, 97, 103)
	Medium	
	Low	
Fear	High	(72, 103)
	Medium	(18, 97)
	Low	(51, 54)
Happiness	**High**	(16, 51, 54)
	Medium	(18, 97)
	Low	
Sadness	High	
	Medium	
	Low	(18, 24, 51, 54, 72, 97)
Tenderness	High	
	Medium	
	Low	(24)
	Formant 1 (*M*)	
Anger	**High**	(51, 54, 62, 79, 100, 103)
	Medium	
	Low	
Fear	High	(87)
	Medium	
	Low	(51, 54, 79)
Happiness	**High**	(16, 52, 54, 87, 100)
	Medium	(51)
	Low	
Sadness	High	(100)
	Medium	
	Low	(51, 52, 54, 62, 79)
Tenderness	High	
	Medium	
	Low	
	Formant 1 (bandwidth)	
Anger	**Narrow**	(38, 51, 100, 103)
	Wide	
Fear	Narrow	
	Wide	(38, 51)
Happiness	**Narrow**	(38, 100)
	Wide	(51)
Sadness	Narrow	
	Wide	(38, 51, 100)
Tenderness	Narrow	
	Wide	

Table 8 (*continued*)

Emotion	Category	Vocal expression studies
	Jitter	
Anger	**High**	(9, 38, 50, 51, 97, 101)
	Low	(56)
Fear	High	(51, 56, 102, 103)
	Low	(50, 67, 78, 97)
Happiness	**High**	(50, 51, 68, 97, 101)
	Low	(20, 56, 67)
Sadness	High	(56)
	Low	(20, 50, 51, 97, 101)
Tenderness	High	
	Low	
	Glottal waveform	
Anger	**Steep**	(16, 22, 38, 50, 61, 72)
	Rounded	
Fear	Steep	(50, 72)
	Rounded	(16, 18, 38, 56)
Happiness	**Steep**	(50, 56)
	Rounded	
Sadness	Steep	
	Rounded	(38, 50, 56, 61)
Tenderness	Steep	
	Rounded	

Note. Numbers within parantheses refer to studies as numbered in Table 2. Text in bold indicates the most frequent finding for respective acoustic cue.

based on the most extensive review to date, strongly suggest—contrary to some previous reviews (e.g., Russell, Bachorowski, & Fernández-Dols, 2003)—that there are emotion-specific patterns of acoustic cues that can be used to communicate discrete emotions in both vocal and musical expression of emotion.

Theoretical Accounts

Accounting for cross-modal similarities. Similarities between vocal expression and music performance in terms of the acoustic cues used to express specific emotions could, on a superficial level, be interpreted in five ways. First, one could argue that these parallels are merely coincidental—a matter of sheer chance. However, because we have discovered a large number of similarities regarding many different aspects, this interpretation seems far fetched. Second, one might argue that the obtained similarities are due to some third variable—for instance, that both vocal expression and music performance are based on principles of body language. However, that vocal expression and music performance share many characteristics that are unique to acoustic signals (e.g., timbre) renders this explanation less than optimal. Furthermore, an account in terms of body language is less parsimonious than Spencer's law. Why evoke an explanation through a different perceptual modality when there is an explanation within the same modality? Vocal expression of emotions mainly reflects physiological responses associated with specific emotions that have a direct and differentiated impact on the voice organs. Third, one could argue that speakers base their vocal expressions of emotions on how performers express emotions in music. To support this hypothesis one would have to demonstrate that music performers' use of the code logically precedes its use in vocal expression.

However, given phylogenetic continuity of vocal expression of emotion, involving subcortical parts of the brain that humans share with other social mammals (Panksepp, 2000), and that music seems to involve specialized neural networks that are more recent and require cortical mediation (e.g., Peretz, 2001), this hypothesis is implausible. Fourth, one could argue, as indeed some authors have, that both channels evolved in parallel without one preceding the other. However, this argument is similarly inconsistent with neuropsychological results that suggest that those parts of the brain that are concerned with vocal expressions of emotions are probably phylogenetically older than those parts concerned with the processing of musical structures. Finally, one could argue that musicians communicate emotions to listeners on the basis of the principles of vocal expression of emotion. This is the explanation that is advocated here. Human vocal expression of emotion is organized and initiated by evolved affect programs that are also present in nonhuman primates. Hence, vocal expression is the model on which musical expression is based rather than the other way around, as postulated by Spencer's law. This evolutionary perspective is consistent with the present findings that (a) vocal expression of emotions is cross-culturally accurate and (b) decoding of vocal expression of emotions develops early in ontogeny. However, it is crucial to note that our argument applies only to the nonverbal aspects of vocal communication. In our estimation, it is likely that vocal expression of emotions developed first and that music performance developed concurrently with speech (Brown, 2000) or even prior to speech (Darwin, 1872/1998; Rousseau, 1761/1986).

Accounting for inconsistency in code usage. In a previous review of vocal expression published in this journal, Scherer

Table 9
Patterns of Acoustic Cues Used to Express Emotions Specifically in Music Performance Studies

Emotion	Category	Music performance studies		
		Articulation (M; d_{io}/d_{ii})		
Anger	**Staccato**	(2, 7, 8, 9, 10, 13, 14, 19, 23, 29)		
	Legato	(16, 18, 20, 22, 25)		
Fear	**Staccato**	(2, 7, 13, 18, 19, 20, 22, 23, 25)		
	Legato	(16)		
Happiness	**Staccato**	(1, 7, 8, 13, 14, 16, 18, 19, 20, 22, 23, 29, 35)		
	Legato	(9, 25)		
Sadness	Staccato			
	Legato	(2, 7, 8, 9, 13, 16, 18, 19, 20, 21, 22, 23, 25, 29)		
Tenderness	Staccato			
	Legato	(1, 2, 7, 8, 12, 13, 14, 16, 19)		
		Articulation (SD; d_{io}/d_{ii})		
Anger	Large			
	Medium	(14, 18, 20, 22, 23)		
	Small			
Fear	**Large**	(18, 20, 22, 23)		
	Medium			
	Small			
Happiness	**Large**	(10, 14, 18, 20, 22, 23)		
	Medium			
	Small			
Sadness	Large			
	Medium			
	Small	(18, 20, 22, 23)		
Tenderness	Large			
	Medium			
	Small	(14)		
		Vibrato (magnitude/rate)		
Anger	**Large**	(13, 15, 16, 24, 26, 30, 32, 35)	**Fast**	(24)
	Small		Slow	
Fear	Large	(30, 32)	**Fast**	(13, 19, 24)
	Small	(13, 24, 30)	Slow	
Happiness	Large	(26, 32)	**Fast**	(1, 13, 24)
	Small	(13, 24)	Slow	
Sadness	Large	(24)	Fast	
	Small	(26, 30, 32)	**Slow**	(8, 13, 19, 24, 16)
Tenderness	Large		Fast	
	Small	(30)	**Slow**	(1)
		Timing variability		
Anger	Large			
	Medium	(8, 14, 23)		
	Small	(22, 29)		
Fear	**Large**	(2, 8, 10, 13, 18, 22, 23, 27)		
	Medium			
	Small			
Happiness	Large			
	Medium	(27)		
	Small	(1, 8, 13, 14, 22, 23, 29)		
Sadness	Large	(2, 13, 16)		
	Medium	(10, 22, 23)		
	Small	(29)		
Tenderness	**Large**	(1, 2, 13)		
	Medium			
	Small	(14)		
		Duration contrasts (between long and short notes)		
Anger	**Sharp**	(13, 14, 16, 27, 28)		
	Soft			
Fear	**Sharp**	(27, 28)		
	Soft			

Table 9 (*continued*)

Emotion	Category	Music performance studies
		Duration contrasts (between long and short notes) (*continued*)
Happiness	**Sharp**	(13, 16, 27)
	Soft	(14, 28)
Sadness	Sharp	
	Soft	(13, 16, 27, 28)
Tenderness	Sharp	
	Soft	(13, 14, 16, 27, 28)
		Singer's formant
Anger	**High**	(31)
	Low	
Fear	High	(40)
	Low	(31)
Happiness	**High**	(31)
	Low	
Sadness	High	
	Low	(31, 40)
Tenderness	High	
	Low	(40)

Note. Numbers within parantheses refer to studies as numbered in Table 3. Text in bold indicates the most frequent finding for respective acoustic cue. d_{io} = duration of time from the onset of tone until its offset; d_{ii} = the duration of time from the onset of a tone until the onset of the next tone.

(1986) observed the apparent paradox that listeners are successful at decoding emotions from vocal expressions despite researchers' having found it difficult to identify acoustic cues that reliably differentiate among emotions. In the present review, which has benefited from additional data collected since Scherer's (1986) review, we have found evidence of emotion-specific patterns of cues. Even so, some inconsistency in code usage remains (see Table 7) and requires explanation. We argue that part of this explanation should be sought in terms of the coding of the communicative process.

Studies of vocal expression and studies of music performance have shown that the relevant cues are coded probabilistically,

Table 10
Comparison of Results for Acoustic Cues in Vocal Expression With Scherer's (1986) Predictions

Acoustic cue	Main finding/prediction by emotion category			
	Anger	Fear	Happiness	Sadness
Speech rate	+/+	+/+	+/+	−/−
Intensity (*M*)	+/+	+/+	+/+	−/−
Intensity (*SD*)	+/+	+/+	+/+	−/−
F0 (*M*)	+/	+/+	+/+	−/±
F0 (*SD*)	+/+	−/+	+/+	−/−
F0 contour[a]	+/=	+/+	+/+	−/−
High-frequency energy	+/+	+/+	+/±	−/±
Formant 1 (*M*)	+/+	−/+	+/−	−/−

Note. Only the direction of the effect (positive [+] vs. negative [−]) is indicated. No predictions were made by Scherer (1986) for the tenderness category or for mean fundamental frequency (F0) in the anger category. ± = predictions in opposing directions.
[a] For F0 contour, a plus sign indicates an upward contour, a minus sign indicates a downward contour, and an equal sign indicates no change.

continuously, and iconically (e.g., Juslin, 1997b; Scherer, 1982). Furthermore, there are intercorrelations between the cues, and these correlations are of about the same magnitude in both channels (Banse & Scherer, 1996; Juslin, 2000; Juslin & Laukka, 2001). These features of the coding could explain many characteristics of the communicative process in both vocal expression and music performance. To capture these characteristics, one might benefit from consideration of Brunswik's (1956) conceptual framework (Hammond & Stewart, 2001). Specifically, it has been suggested that Brunswik's lens model may be useful to describe the communicative process in vocal expression (Scherer, 1982) and music performance (Juslin, 1995, 2000). The lens model was originally intended as a model of visual perception, capturing relations between an organism and distal cues.[12] However, it was later used mainly in studies of human judgment.

Although Brunswik's (1956) lens model failed as a full-fledged model of visual perception, it seems highly appropriate for describing communication of emotion. Specifically, the lens model can be used to illustrate how encoders express specific emotions by using a large set of cues (e.g., speed, intensity, timbre) that are probabilistic (i.e., uncertain) though partly redundant. The emotions are recognized by decoders, who use the same cues to decode the emotional expression. The cues are probabilistic in that they are not perfectly reliable indicators of the expressed emotion. Therefore, decoders have to combine many cues for successful communication to occur. This is not simply a matter of pattern matching, however, because the cues contribute in an additive fashion to decoders' judgments. Brunswik's concept of vicarious functioning can be used to capture how decoders use the partly interchangeable cues in flexible ways by occasionally shifting

_____

[12] In fact, Brunswik (1956) applied the model to facial expression of emotion, among other things (pp. 111–113).

Table 11
Summary of Cross-Modal Patterns of Acoustic Cues for Discrete Emotions

Emotion	Acoustic cues (vocal expression/music performance)
Anger	Fast speech rate/tempo, high voice intensity/sound level, much voice intensity/sound level variability, much high-frequency energy, high F0/pitch level, much F0/pitch variability, rising F0/pitch contour, fast voice onsets/tone attacks, and microstructural irregularity
Fear	Fast speech rate/tempo, low voice intensity/sound level (except in panic fear), much voice intensity/sound level variability, little high-frequency energy, high F0/pitch level, little F0/pitch variability, rising F0/pitch contour, and a lot of microstructural irregularity
Happiness	Fast speech rate/tempo, medium–high voice intensity/sound level, medium high-frequency energy, high F0/pitch level, much F0/pitch variability, rising F0/pitch contour, fast voice onsets/tone attacks, and very little microstructural regularity
Sadness	Slow speech rate/tempo, low voice intensity/sound level, little voice intensity/sound level variability, little high-frequency energy, low F0/pitch level, little F0/pitch variability, falling F0/pitch contour, slow voice onsets/tone attacks, and microstructural irregularity
Tenderness	Slow speech rate/tempo, low voice intensity/sound level, little voice intensity/sound level variability, little high-frequency energy, low F0/pitch level, little F0/pitch variability, falling F0/pitch contours, slow voice onsets/tone attacks, and microstructural regularity

Note. F0 = fundamental frequency.

from a cue that is unavailable to one that is available (Juslin, 2001a).

The findings reviewed in this article are consistent with Brunswik's (1956) lens model. First, it is clear that the cues are probabilistically related only to encoding and decoding. The probabilistic nature of the cues reflects (a) individual differences between encoders, (b) structural constraints of the verbal or musical material used, and (c) that the same cue can be used in the same way in more than one expression. For instance, fast speed can be used in both happiness and anger, and therefore speech rate is not a perfect indicator of either emotion. Second, evidence confirms that cues contribute in an additive fashion to listeners' judgments, as shown by a general lack of cue interactions (Juslin, 1997c; Ladd, Silverman, Tolkmitt, Bergmann, & Scherer, 1985; Scherer & Oshinsky, 1977), and that emotions can be communicated successfully on different instruments that provide relatively different, though partly interchangeable, acoustic cues to the performer's disposal. (If a performer cannot vary the timbre to express anger, he or she compensates this by varying the loudness even more.) Each cue is neither necessary nor sufficient, but the larger the number of cues used, the more reliable the communication (Juslin, 2000). Third, a Brunswikian conceptualization of the communicative process in terms of separate cues that are integrated—as opposed to a "Gibsonian" (Gibson, 1979) conceptualization, which conceives of the process in terms of holistic higher order-variables—is supported by studies on the physiology of listening. Handel (1991) noted that speech and music seem to involve similar perceptual mechanisms. The auditory pathways involve different neural representations for various aspects of the acoustic signal (e.g., timing, frequency), which are kept separate until later stages of analysis. Perception of both speech and music requires the integration of these different representations, as implied by the lens model. Fourth, and as noted above, there is strong evidence of intercorrelations (i.e., redundancy) among acoustic cues. The redundancy between cues largely reflects the sound production mechanisms of the voice and of musical instruments. For instance, an increase in subglottal pressure (i.e., the air pressure in the lungs driving the speech) increases not only the intensity but also the F0 to some degree. Similarly, a harder string attack produces a tone

that is both louder and sharper in timbre (the occurrence of these effects partly reflects fundamental physical principles, such as nonlinear excitation; Wolfe, 2002).

The coding captured by Brunswik's (1956) lens model has one particularly important implication: Because the acoustic cues are intercorrelated to some degree, more than one way of using the cues might lead to a similarly high level of decoding accuracy (e.g., Dawes & Corrigan, 1974; Juslin, 2000). The lens model might explain why we found accurate communication of emotions in vocal expression and music performance (see findings of meta-analysis in Results section) despite considerable inconsistency in code usage (see Tables 7–9); multiple cues that are partly redundant yield a robust communicative system that is forgiving of deviations from optimal code usage. Performers are thus able to communicate emotions to listeners without having to compromise their unique playing styles (Juslin, 2000). Similarly, it may be expected that different actors communicate emotions successfully in different ways, thereby avoiding stereotypical portrayals of emotions in theater. However, this robustness comes with a price. The redundancy of the cues means that the same information is conveyed by many cues. This limits the information capacity of the channel (Juslin, 1998; see also Shannon & Weaver, 1949). This may explain why encoders are able to communicate broad emotion categories but not finer nuances within the categories (e.g., Dowling & Harwood, 1986, chap. 8; Greasley, Sherrard, & Waterman, 2000; Juslin, 1997a; L. Kaiser, 1962; London, 2002). A communication system of this type shows "compromise and a falling short of precision, but also the relative infrequency of drastic error" (Brunswik, 1956, p. 145). An evolutionary perspective may explain this characteristic: It is ultimately more important to avoid making serious mistakes (e.g., mistaking anger for sadness) than to be able to make more subtle discriminations between emotions (e.g., detecting different kinds of anger). Redundancy in the coding helps to counteract the degradation of acoustic signals during transmission that occur in natural environments because of factors such as attenuation and reverberation (Wiley & Richards, 1978).

Accounting for induction of emotions. This review has revealed similarities in the acoustic cues used to communicate emotions in vocal expression and music performance. Can these find-

ings also explain induction of emotions in listeners? We propose that listeners can become "moved" by music performances through a process of emotional contagion (Hatfield, Cacioppo, & Rapson, 1994). Evidence suggests that people easily "catch" the emotions of others when seeing their facial expressions or hearing their vocal expressions (see Neumann & Strack, 2000). If performances of music express emotions in ways similar to how voices express emotions, it follows that people could get aroused by the voicelike aspect of music.[13] Evidence that individuals do react emotionally to music as they do to vocal expressions of emotion comes from investigations using facial electromyography and self-reports to measure emotion (Hietanen, Surakka, & Linnankoski, 1998; Lundqvist, Carlsson, & Hilmersson, 2000; Neumann & Strack, 2000; Witvliet & Vrana, 1996; Witvliet, Vrana, & Webb-Talmadge, 1998).

Some authors, however, have argued that music performances do not sound very much like vocal expressions, at least superficially (Budd, 1985, chap. 7). Why, then, should individuals respond to music performances as if they were vocal expressions? One explanation is that expressions of emotion are processed by domain-specific and autonomous "modules" of the brain (Fodor, 1983), which react to certain acoustic features in the stimulus. The emotion perception modules do not recognize the difference between vocal expressions and other acoustic expressions and therefore react in much the same way (e.g., registering anger) as long as certain cues (e.g., high speed, loud dynamics, rough timbre) are present in the stimulus. The modular view of information processing has been the subject of much debate in recent years (cf. Coltheart, 1999; Geary & Huffman, 2002; Öhman & Mineka, 2001; Pinker, 1997), although even some of its most ardent critics have admitted that special-purpose modules may indeed exist at the subcortical level of the brain, where much of the processing of emotion occurs (Panksepp & Panksepp, 2000).

Although a modular theory of emotion perception in music remains to be fully investigated, limited support for such a theory in terms of Fodor's (1983) proposed characteristics of modules (see also Coltheart, 1999) comes from evidence (a) of brain dissociations between judgments of musical emotion and of musical structure (Peretz, Gagnon, & Bouchard, 1998; modules are domain-specific), (b) that judgments of musical emotions are quick (Peretz et al., 1998; modules are fast), (c) that the ability to decode emotions from music develops early (Cunningham & Sterling, 1988; modules are innately specified), (d) that processing in the perception of emotional expression is primarily implicit (Niedenthal & Showers, 1991; modules are autonomous), (e) that it is impossible to relearn how to associate expressive forms with emotions (Clynes, 1977, p. 45; modules are hard-wired), (f) that emotion induction through music is possible even if listeners do not attend to the music (Västfjäll, 2002; modules are automatic), and (g) that individuals react to music performances as if they were expressions of emotion (Witvliet & Vrana, 1996) despite knowing that music does not literally have emotions to express (modules are information capsulated).

One problem with the present approach is that it seems to ignore the unique value of music (Budd, 1985, chap. 7). As noted by several authors, music is not only a tool for communicating emotion. Therefore, we must reach beyond this notion and explain why people listen to music specifically, rather than to just any expression of emotion. One way around this problem would be to identify

ways in which musical expression is special (apart from occurring in music). Juslin (in press) argued that what makes a particular music performance of, say, the violin, so expressive is the fact that it sounds a lot like the human voice while going far beyond what the human voice can do (e.g., in terms of speed, pitch range, and timbre). Consequently, we speculate that many musical instruments are processed by brain modules as superexpressive voices. For instance, if human speech is perceived as angry when it has fast rate, loud intensity, and harsh timbre, a musical instrument might sound extremely angry in virtue of its even higher speed, louder intensity, and harsher timbre. The "attention" of the emotion-perception module is gripped by the music's voicelike nature, and the individual becomes aroused by the extreme turns taken by this voice. The emotions evoked in listeners may not necessarily be the same as those expressed and perceived but could be empathic or complementary (Juslin & Zentner, 2002). We admit that these ideas are speculative, but we think that they merit further study given that similarities between vocal expression and music performance have been obtained. We emphasize that this is only one of many possible sources of musical emotions (Juslin, 2003; Scherer & Zentner, 2001).

Problems and Directions for Future Research

In this section, we identify important problems and suggest directions for future research. First, given the large individual differences in encoding accuracy and code usage, researchers must ensure that reasonably large samples of encoders are used. In particular, researchers must avoid studying only one encoder, because doing so may cause serious threats to the external validity of the study. For instance, it may be impossible to know whether the obtained findings for a particular emotion can be generalized to other encoders.

Second, researchers should pay closer attention to the precise contents to be communicated, preferably basing choices of emotion labels on theoretical grounds (Juslin, 1997b; Scherer, 1986). Studies of music performance, in particular, have frequently included emotion labels without any consideration of what contents are theoretically or musically plausible. The results, both in terms of communication accuracy and consistency of code usage, are likely to differ greatly, depending on the emotion labels used (Juslin, 1997b). This point is brought home by the low accuracy reported in studies that used more abstract labels, such as *deep* or *sophisticated* (Senju & Ohgushi, 1987). Moreover, the use of more well-differentiated emotion labels, in terms of both quantity (Juslin & Laukka, 2001) and quality (Banse & Scherer, 1996) of emotion, could help to reduce some of the inconsistency in empirical findings.

Third, we recommend that researchers study encoding and decoding in a combined fashion such that the two aspects may be related. Only if encoding and decoding processes are analyzed in

---

[13] Many authors have proposed that vocal and musical expression of emotion is especially effective in causing emotional contagion (Eibl-Eibesfeldt, 1989, p. 691; Lewis, 2000, p. 270). One possible explanation may be that hearing is the perceptual modality that develops first. In fact, because hearing is functional even prior to birth, some relations between acoustic patterns and emotional states may reflect prenatal experiences (Mastropieri & Turkewitz, 1999, p. 205).

combination can a more complete understanding of the communicative process be reached. This is a prerequisite if one intends to improve communication (Juslin & Laukka, 2000). Brunswik's (1956) lens model and the accompanying lens model equation (Hursch, Hammond, & Hursch, 1964) could be useful tools in attempts to relate encoding to decoding in vocal expression (Scherer, 1978, 1982) and music performance (Juslin, 1995, 2000). The lens model shows that the success of any communicative process depends equally on the encoder and the decoder. Uncertainty is an unavoidable aspect of this process, and multiple regression analysis may be suitable for capturing the uncertain relationships among encoders, cues, and decoders (e.g., Hargreaves, Starkweather, & Blacker, 1965; Juslin, 2000; Roessler & Lester, 1976; Scherer & Oshinsky, 1977).

Fourth, much remains to be done concerning the measurement of acoustic cues. There is an abundance of studies that analyze only a few cues, but there is an urgent need for studies that try to describe the complete set of cues. If not all relevant cues are captured, researchers run the risk of leaving out important aspects of the code. Estimates of the relative importance of cues are then likely to be grossly misleading. A challenge for future research is to go beyond the classic cues (pitch, speed, intensity) and try to analyze more subtle cues, such as continuously varying patterns of speed and dynamics. For assistance, researchers may use computer programs that allow them to extract characteristic timing patterns from emotion portrayals. These patterns might be used in synthesized sound sequences to examine their effects on listeners' judgments (Juslin & Madison, 1999). Finally, researchers should take greater care in reporting the data for all acoustic cues and emotions. Many articles provide only partial reports of the data. This problem prevented us from conducting a meta-analysis of the results regarding code usage. By more carefully reporting data, researchers could contribute to the development of more precise quantitative predictions.

Fifth, studies of vocal expression and music performance have primarily been conducted in tightly controlled laboratory settings. Far less is known about these phenomena as they occur in more ecologically valid settings. In studies of vocal expression, a crucial question concerns how similar emotion portrayals are to natural expressions (Bachorowski, 1999). Unfortunately, the number of studies that have used natural speech is too small to permit definitive conclusions. As regards music, certain authors have cautioned that performances recorded under experimental conditions may lead to different results than performances made under natural conditions, such as concerts (Rapoport, 1996). Again, relevant evidence is still lacking. To conduct ecologically valid studies without sacrificing internal validity represents a challenge for future research.

Sixth, findings from analyses of acoustic cues should be evaluated in listening experiments using synthesized sound sequences to test specific hypotheses (Table 11). Because cues in vocal expression and music performance are probabilistic and intercorrelated to some degree, only by using synthesized sound sequences that are systematically manipulated in a factorial design can one establish that a given cue really has predictable effects on listeners' judgments of expression. Synthesized sound sequences may be regarded as computational models, which demonstrate the validity of proposed hypotheses by showing that they really work (Juslin, 1996; Juslin, 1997c; Juslin, Friberg, & Bresin, 2002; Murray &

Arnott, 1993, 1995). It should be noted that although encoders may not use a particular cue (e.g., vibrato, jitter) in a consistent fashion, the cue might still be reliably associated with decoders' emotion judgments, as indicated by listening tests with synthesized stimuli. The opposite may also be true: encoders may use a given cue in a consistent fashion, but decoders may fail to use this cue. This highlights the importance of studying both encoding and decoding aspects of the communicative process (see also Buck, 1984, chap. 5–7).

Seventh, a greater variety of verbal or musical materials should be used in future research to maximize its generalizability. Researchers have often assumed that the encoding proceeds more or less independently of the material, but this assumption has been questioned (Cosmides, 1983; Juslin, 1998; Scherer, 1986). Although some evidence of a dissociation between linguistic stress and emotion has been obtained (McRoberts, Studdert-Kennedy, & Shankweiler, 1995), it seems unlikely that variability in cues (e.g., fundamental frequency, timing) that function linguistically as semantic and syntactic markers in speech (Scherer, 1979) and music performance (Carlson, Friberg, Frydén, Granström, & Sundberg, 1989) leaves the emotional expression completely unaffected. On the contrary, because most studies have included only one set of verbal or musical material, it is possible that inconsistency in previous data reflects interactions between materials and acoustic cues. Future research would arguably benefit from a closer study of such interactions (Cowie et al., 2001; Juslin, 1998, p. 50).

Eighth, the use of forced-choice formats has been criticized on the grounds that listeners are provided with only a small number of response alternatives to choose from (Ekman, 1994; Izard, 1994; Russell, 1994). It may be argued that listeners manage the task by forming exclusion rules or guessing, without thinking that any of the response alternatives are appropriate to describe the expression (e.g., Frick, 1985). Those studies that have used free labeling of emotions, rather than forced-choice formats, indicate that communication is still possible, though the accuracy is slightly lower (e.g., Juslin, 1997a; L. Kaiser, 1962). Juslin (1997a) suggested that what can be communicated reliably are the basic emotion categories but not specific nuances within these categories. It is desirable to use a wider variety of response formats in future research (see also Greasley et al., 2000).

Finally, the two domains could benefit from studies of the neurophysiological substrates of the decoding process (Adolphs, Damasio, & Tranel, 2002). For example, it would be interesting to explore whether the same neurological resources are used in decoding of emotions from vocal expression and music performance (Peretz, 2001). It must be noted that if the neural circuitry used in decoding of emotion from vocal expression is also involved in decoding of emotion from music, this should primarily apply to those aspects of music's expressiveness that are common to speech and music performance; that is, cues like speed, intensity, and timbre. However, music's expressiveness does not derive solely from those cues but also from intrinsic sources of emotion (see Sloboda & Juslin, 2001) having to do with the structure of the piece of music (e.g., harmony). Thus, it would not be surprising if perception of emotion in music involved neural substrates over and above those involved in perception of vocal expression. Preliminary evidence that perception of emotion in music performance involves many of the same brain areas as perception of emotion in

vocal expression was reported by Nair, Large, Steinberg, and Kelso (2002).

Concluding Remarks

Research on communication of emotions might lead to a number of important applications. First, research on vocal cues might be used to develop instruments for the diagnosis of different psychiatric conditions, such as depression and schizophrenia (S. Kaiser & Scherer, 1998). Second, results regarding code usage might be used in teaching of rhetoric. Pathos, or emotional appeal, is considered an important means of persuasion (e.g., Lee, 1939), and this article offers detailed information about the practical means to convey specific emotions to an audience. Third, recent research on communication of emotion might be used by music teachers to enhance performers' expressivity (Juslin, Friberg, Schoonderwaldt, & Karlsson, in press; Juslin & Persson, 2002). Fourth, communication of emotions can be trained in music therapy. Proficiency in emotional communication is part of the emotional intelligence (Salovey & Mayer, 1990) that most people take for granted but that certain individuals lack. Music provides a way of training encoding and decoding of emotions in a fairly nonthreatening situation (for reviews, see Saperston, West, & Wigram, 1995). Finally, research on vocal communication of emotion has implications for human–computer interaction, especially automatic recognition of emotion and synthesis of emotional speech (Cowie et al., 2001; Murray & Arnott, 1995; Schröder, 2001).

In conclusion, a number of authors have speculated about an intimate relationship between vocal expression and music performance regarding communication of emotions. This article has reached beyond the speculative stage and established many similarities among the two channels in terms of decoding accuracy, code usage, development, and coding. It is our strong belief that continued cross-modal research will provide further insights into the expressive aspects of vocal expression and music performance—insights that would be difficult to obtain from studying the two domains in separation. In particular, we predict that future research will confirm that music performers communicate emotions to listeners by exploiting an acoustic code that derives from innate brain programs for vocal expression of emotions. In this sense, at least, music may really be a form of heightened speech that transforms feelings into "audible landscape."

References

References marked with an asterisk indicate studies included in the meta-analysis.

Abelin, Å., & Allwood, J. (2000). Cross linguistic interpretation of emotional prosody. In R. Cowie, E. Douglas-Cowie, & M. Schröder (Eds.), *Proceedings of the ISCA workshop on speech and emotion* [CD-ROM]. Belfast, Ireland: International Speech Communication Association.

Adachi, M., & Trehub, S. E. (2000). Decoding the expressive intentions in children's songs. *Music Perception, 18,* 213–224.

Adolphs, R., Damasio, H., & Tranel, D. (2002). Neural systems for recognition of emotional prosody: A 3-D lesion study. *Emotion, 2,* 23–51.

Albas, D. C., McCluskey, K. W., & Albas, C. A. (1976). Perception of the emotional content of speech: A comparison of two Canadian groups. *Journal of Cross-Cultural Psychology, 7,* 481–490.

Alpert, M., Pouget, E. R., & Silva, R. R. (2001). Reflections of depression in acoustic measures of the patient's speech. *Journal of Affective Disorders, 66,* 59–69.

*Al-Watban, A. M. (1998). *Psychoacoustic analysis of intonation as a carrier of emotion in Arabic and English.* Unpublished doctoral dissertation, Ball State University, Muncie, IN.

*Anolli, L., & Ciceri, R. (1997). *La voce delle emozioni* [The voice of emotions]. Milan: FrancoAngeli.

Apple, W., & Hecht, K. (1982). Speaking emotionally: The relation between verbal and vocal communication of affect. *Journal of Personality and Social Psychology, 42,* 864–875.

Arcos, J. L., Cañamero, D., & López de Mántaras, R. (1999). Affect-driven CBR to generate expressive music. *Lecture Notes in Artificial Intelligence, 1650,* 1–13.

Baars, G., & Gabrielsson, A. (1997). Emotional expression in singing: A case study. In A. Gabrielsson (Ed.), *Proceedings of the Third Triennial ESCOM Conference* (pp. 479–483). Uppsala, Sweden: Uppsala University.

Bachorowski, J.-A. (1999). Vocal expression and perception of emotion. *Current Directions in Psychological Science, 8,* 53–57.

Bachorowski, J.-A., & Owren, M. J. (1995). Vocal expression of emotion: Acoustical properties of speech are associated with emotional intensity and context. *Psychological Science, 6,* 219–224.

Balkwill, L.-L., & Thompson, W. F. (1999). A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music Perception, 17,* 43–64.

Baltaxe, C. A. M. (1991). Vocal communication of affect and its perception in three- to four-year-old children. *Perceptual and Motor Skills, 72,* 1187–1202.

*Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology, 70,* 614–636.

*Baroni, M., Caterina, R., Regazzi, F., & Zanarini, G. (1997). Emotional aspects of singing voice. In A. Gabrielsson (Ed.), *Proceedings of the Third Triennial ESCOM Conference* (pp. 484–489). Uppsala, Sweden: Uppsala University.

Baroni, M., & Finarelli, L. (1994). Emotions in spoken language and in vocal music. In I. Deliège (Ed.), *Proceedings of the Third International Conference for Music Perception and Cognition* (pp. 343–345). Liége, Belgium: University of Liége.

Becker, J. (2001). Anthropological perspectives on music and emotion. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 135–160). New York: Oxford University Press.

Behrens, G. A., & Green, S. B. (1993). The ability to identify emotional content of solo improvisations performed vocally and on three different instruments. *Psychology of Music, 21,* 20–33.

Bekoff, M. (2000). Animal emotions: Exploring passionate natures. *BioScience, 50,* 861–870.

Bengtsson, I., & Gabrielsson, A. (1980). Methods for analyzing performances of musical rhythm. *Scandinavian Journal of Psychology, 21,* 257–268.

Bergmann, G., Goldbeck, T., & Scherer, K. R. (1988). Emotionale Eindruckswirkung von prosodischen Sprechmerkmalen [The effects of prosody on emotion inference]. *Zeitschrift für Experimentelle und Angewandte Psychologie, 35,* 167–200.

Besson, M., & Friederici, A. D. (1998). Language and music: A comparative view. *Music Perception, 16,* 1–9.

Bloch, S., Orthous, P., & Santibañez, G. (1987). Effector patterns of basic emotions: A psychophysiological method for training actors. *Journal of Social Biology and Structure, 10,* 1–19.

Bonebright, T. L. (1996). *Vocal affect expression: A comparison of multidimensional scaling solutions for paired comparisons and computer sorting tasks using perceptual and acoustic measures.* Unpublished doctoral dissertation, University of Nebraska, Lincoln.

*Bonebright, T. L., Thompson, J. L., & Leger, D. W. (1996). Gender

stereotypes in the expression and perception of vocal affect. *Sex Roles, 34,* 429–445.

Bonner, M. R. (1943). Changes in the speech pattern under emotional tension. *American Journal of Psychology, 56,* 262–273.

Booth, R. J., & Pennebaker, J. W. (2000). Emotions and immunity. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 558–570). New York: Guilford Press.

Borden, G. J., Harris, K. S., & Raphael, L. J. (1994). *Speech science primer: Physiology, acoustics and perception of speech* (3rd ed.). Baltimore: Williams & Wilkins.

Breitenstein, C., Van Lancker, D., & Daum, I. (2001). The contribution of speech rate and pitch variation to the perception of vocal emotions in a German and an American sample. *Cognition & Emotion, 15,* 57–79.

Bresin, R., & Friberg, A. (2000). Emotional coloring of computer-controlled music performance. *Computer Music Journal, 24,* 44–63.

Breznitz, Z. (1992). Verbal indicators of depression. *Journal of General Psychology, 119,* 351–363.

*Brighetti, G., Ladavas, E., & Ricci Bitti, P. E. (1980). Recognition of emotion expressed through voice. *Italian Journal of Psychology, 7,* 121–127.

Brosgole, L., & Weisman, J. (1995). Mood recognition across the ages. *International Journal of Neuroscience, 82,* 169–189.

Brown, S. (2000). The "musilanguage" model of music evolution. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 271–300). Cambridge, MA: MIT Press.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments.* Berkeley: University of California Press.

Buck, R. (1984). *The communication of emotion.* New York: Guilford Press.

Budd, M. (1985). *Music and the emotions. The philosophical theories.* London: Routledge.

*Bunt, L., & Pavlicevic, M. (2001). Music and emotion: Perspectives from music therapy. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 181–201). New York: Oxford University Press.

*Burkhardt, F. (2001). *Simulation emotionaler Sprechweise mit Sprachsyntheseverfahren* [Simulation of emotional speech by means of speech synthesis]. Doctoral dissertation, Technische Universität Berlin, Berlin, Germany.

*Burns, K. L., & Beier, E. G. (1973). Significance of vocal and visual channels in the decoding of emotional meaning. *Journal of Communication, 23,* 118–130.

Buss, D. M. (1995). Evolutionary psychology: A new paradigm for psychological science. *Psychological Inquiry, 6,* 1–30.

Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M., & Ito, T. A. (2000). The psychophysiology of emotion. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 173–191). New York: Guilford Press.

Cahn, J. E. (1990). The generation of affect in synthesized speech. *Journal of the American Voice I/O Society, 8,* 1–19.

Camras, L. A., & Allison, K. (1985). Children's understanding of emotional facial expressions and verbal labels. *Journal of Nonverbal Behavior, 9,* 89–94.

Canazza, S., & Orio, N. (1999). The communication of emotions in jazz music: A study on piano and saxophone performances. In M. O. Belardinelli & C. Fiorelli (Eds.), *Musical behavior and cognition* (pp. 261–276). Rome: Edizioni Scientifiche Magi.

Carlson, R., Friberg, A., Frydén, L., Granström, B., & Sundberg, J. (1989). Speech and music performance: Parallels and contrasts. *Contemporary Music Review, 4,* 389–402.

Carlson, R., Granström, B., & Nord, L. (1992). Experiments with emotive speech: Acted utterances and synthesized replicas. In J. J. Ohala, T. M. Nearey, B. L. Derwing, M. M. Hodge, & G. E. Wiebe (Eds.), *Proceedings of the Second International Conference on Spoken Language Processing* (pp. 671–674). Edmonton, Alberta, Canada: University of Alberta.

Carterette, E. C., & Kendall, R. A. (1999). Comparative music perception and cognition. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 725–791). San Diego, CA: Academic Press.

*Chung, S.-J. (2000). *L'expression et la perception de l'émotion dans la parole spontanée: Évidences du coréen et de l'anglais* [Expression and perception of emotion extracted from spontaneous speech in Korean and English]. Unpublished doctoral dissertation, Université de la Sorbonne Nouvelle, Paris.

Clynes, M. (1977). *Sentics: The touch of emotions.* New York: Doubleday.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Coltheart, M. (1999). Modularity and cognition. *Trends in Cognitive Science, 3,* 115–119.

Conway, M. A., & Bekerian, D. A. (1987). Situational knowledge and emotions. *Cognition & Emotion, 1,* 145–191.

Cordes, I. (2000). Communicability of emotion through music rooted in early human vocal patterns. In C. Woods, G. Luck, R. Brochard, F. Seddon, & J. A. Sloboda (Eds.), *Proceedings of the Sixth International Conference on Music Perception and Cognition, August 2000* [CD-ROM]. Keele, England: Keele University.

Cosmides, L. (1983). Invariances in the acoustic expression of emotion during speech. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 864–881.

Cosmides, L., & Tooby, J. (1994). Beyond intuition and instinct blindness: Toward an evolutionarily rigorous cognitive science. *Cognition, 50,* 41–77.

Cosmides, L., & Tooby, J. (2000). Evolutionary psychology and the emotions. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 91–115). New York: Guilford Press.

Costanzo, F. S., Markel, N. N., & Costanzo, P. R. (1969). Voice quality profile and perceived emotion. *Journal of Counseling Psychology, 16,* 267–270.

Cowie, R., Douglas-Cowie, E., & Schröder, M. (Eds.). (2000). *Proceedings of the ISCA workshop on speech and emotion* [CD-ROM]. Belfast, Ireland: International Speech Communication Association.

Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., & Taylor, J. G. (2001). Emotion recognition in human-computer interaction. *IEEE Signal Processing Magazine, 18,* 32–80.

Cummings, K. E., & Clements, M. A. (1995). Analysis of the glottal excitation of emotionally styled and stressed speech. *Journal of the Acoustical Society of America, 98,* 88–98.

Cunningham, J. G., & Sterling, R. S. (1988). Developmental changes in the understanding of affective meaning in music. *Motivation and Emotion, 12,* 399–413.

Dalla Bella, S., Peretz, I., Rousseau, L., & Gosselin, N. (2001). A developmental study of the affective value of tempo and mode in music. *Cognition, 80,* B1–B10.

Damasio, A. R., Grabowski, T. J., Bechara, A., Damasio, H., Ponto, L. L. B., Parvizi, J., & Hichwa, R. D. (2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nature Neuroscience, 3,* 1049–1056.

Darwin, C. (1998). *The expression of the emotions in man and animals* (3rd ed.). London: Harper-Collins. (Original work published 1872)

Davies, J. B. (1978). *The psychology of music.* London: Hutchinson.

Davies, S. (2001). Philosophical perspectives on music's expressiveness. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 23–44). New York: Oxford University Press.

*Davitz, J. R. (1964a). Auditory correlates of vocal expressions of emotional meanings. In J. R. Davitz (Ed.), *The communication of emotional meaning* (pp. 101–112). New York: McGraw-Hill.

Davitz, J. R. (1964b). Personality, perceptual, and cognitive correlates of

emotional sensitivity. In J. R. Davitz (Ed.), *The communication of emotional meaning* (pp. 57–68). New York: McGraw-Hill.

Davitz, J. R. (1964c). A review of research concerned with facial and vocal expressions of emotion. In J. R. Davitz (Ed.), *The communication of emotional meaning* (pp. 13–29). New York: McGraw-Hill.

*Davitz, J. R., & Davitz, L. J. (1959). The communication of feelings by content-free speech. *Journal of Communication, 9,* 6–13.

Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin, 81,* 95–106.

Dawes, R. M., & Kramer, E. (1966). A proximity analysis of vocally expressed emotion. *Perceptual and Motor Skills, 22,* 571–574.

de Gelder, B., Teunisse, J. P., & Benson, P. J. (1997). Categorical perception of facial expressions and their internal structure. *Cognition & Emotion, 11,* 1–23.

de Gelder, B., & Vroomen, J. (1996). Categorical perception of emotional speech. *Journal of the Acoustical Society of America, 100,* 2818.

Dimitrovsky, L. (1964). The ability to identify the emotional meaning of vocal expression at successive age levels. In J. R. Davitz (Ed.), *The communication of emotional meaning* (pp. 69–86). New York: McGraw-Hill.

Dolgin, K., & Adelson, E. (1990). Age changes in the ability to interpret affect in sung and instrumentally-presented melodies. *Psychology of Music, 18,* 87–98.

Dowling, W. J., & Harwood, D. L. (1986). *Music cognition.* New York: Academic Press.

Drummond, P. D., & Quah, S. H. (2001). The effect of expressing anger on cardiovascular reactivity and facial blood flow in Chinese and Caucasians. *Psychophysiology, 38,* 190–196.

Dry, A., & Gabrielsson, A. (1997). Emotional expression in guitar band performance. In A. Gabrielsson (Ed.), *Proceedings of the Third Triennial ESCOM Conference* (pp. 475–478). Uppsala, Sweden: Uppsala University.

*Dusenbury, D., & Knower, F. H. (1939). Experimental studies of the symbolism of action and voice. II. A study of the specificity of meaning in abstract tonal symbols. *Quarterly Journal of Speech, 25,* 67–75.

*Ebie, B. D. (1999). *The effects of traditional, vocally modeled, kinesthetic, and audio-visual treatment conditions on male and female middle school vocal music students' abilities to expressively sing melodies.* Unpublished doctoral dissertation, Kent State University, Kent, OH.

Eggebrecht, R. (1983). *Sprachmelodie und musikalische Forschungen im Kulturvergleich* [Speech melody and music research in cross-cultural comparison]. Doctoral dissertation, University of Münich, Münich, Germany.

Eibl-Eibesfeldt, I. (1973). The expressive behaviors of the deaf-and-blind-born. In M. von Cranach & I. Vine (Eds.), *Social communication and movement* (pp. 163–194). New York: Academic Press.

Eibl-Eibesfeldt, I. (1989). *Human ethology.* New York: Aldine.

Ekman, P. (Ed.). (1973). *Darwin and facial expression.* New York: Academic Press.

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion, 6,* 169–200.

Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique. *Psychological Bulletin, 115,* 268–287.

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica, 1,* 49–98.

Ekman, P., Levenson, R. W., & Friesen, W. V. (1983, September 16). Autonomic nervous system activity distinguishes between emotions. *Science, 221,* 1208–1210.

Eldred, S. H., & Price, D. B. (1958). A linguistic evaluation of feeling states in psychotherapy. *Psychiatry, 21,* 115–121.

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin, 128,* 203–235.

Ellgring, H., & Scherer, K. R. (1996). Vocal indicators of mood change in depression. *Journal of Nonverbal Behavior, 20,* 83–110.

Erlewine, M., Bogdanow, V., Woodstra, C., & Koda, C. (Eds.). (1996). *The blues.* San Francisco: Miller-Freeman.

Etcoff, N. L., & Magee, J. J. (1992). Categorical perception of facial expressions. *Cognition, 44,* 227–240.

Fairbanks, G., & Hoaglin, L. W. (1941). An experimental study of the durational characteristics of the voice during the expression of emotion. *Speech Monographs, 8,* 85–90.

*Fairbanks, G., & Provonost, W. (1938, October 21). Vocal pitch during simulated emotion. *Science, 88,* 382–383.

Fairbanks, G., & Provonost, W. (1939). An experimental study of the pitch characteristics of the voice during the expression of emotion. *Speech Monographs, 6,* 87–104.

*Fenster, C. A., Blake, L. K., & Goldstein, A. M. (1977). Accuracy of vocal emotional communications among children and adults and the power of negative emotions. *Journal of Communications Disorders, 10,* 301–314.

Fodor, J. A. (1983). *The modularity of the mind.* Cambridge, MA: MIT Press.

Fónagy, I. (1978). A new method of investigating the perception of prosodic features. *Language and Speech, 21,* 34–49.

Fónagy, I., & Magdics, K. (1963). Emotional patterns in intonation and music. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung, 16,* 293–326.

Frick, R. W. (1985). Communicating emotion: The role of prosodic features. *Psychological Bulletin, 97,* 412–429.

Frick, R. W. (1986). The prosodic expression of anger: Differentiating threat and frustration. *Aggressive Behavior, 12,* 121–128.

Fridlund, A. J., Schwartz, G. E., & Fowler, S. C. (1984). Pattern recognition of self-reported emotional state from multiple-site facial EMG activity during affective imagery. *Psychophysiology, 21,* 622–637.

Friend, M. (2000). Developmental changes in sensitivity to vocal paralanguage. *Developmental Science, 3,* 148–162.

Friend, M., & Farrar, M. J. (1994). A comparison of content-masking procedures for obtaining judgments of discrete affective states. *Journal of the Acoustical Society of America, 96,* 1283–1290.

Fulcher, J. A. (1991). Vocal affect expression as an indicator of affective response. *Behavior Research Methods, Instruments, & Computers, 23,* 306–313.

Gabrielsson, A. (1995). Expressive intention and performance. In R. Steinberg (Ed.), *Music and the mind machine* (pp. 35–47). Heidelberg, Germany: Springer.

Gabrielsson, A. (1999). The performance of music. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 501–602). San Diego, CA: Academic Press.

Gabrielsson, A., & Juslin, P. N. (1996). Emotional expression in music performance: Between the performer's intention and the listener's experience. *Psychology of Music, 24,* 68–91.

Gabrielsson, A., & Juslin, P. N. (2003). Emotional expression in music. In R. J. Davidson, H. H. Goldsmith, & K. R. Scherer (Eds.), *Handbook of affective sciences* (pp. 503–534). New York: Oxford University Press.

Gabrielsson, A., & Lindström, E. (1995). Emotional expression in synthesizer and sentograph performance. *Psychomusicology, 14,* 94–116.

Gårding, E., & Abramson, A. S. (1965). A study of the perception of some American English intonation contours. *Studia Linguistica, 19,* 61–79.

Geary, D. C., & Huffman, K. J. (2002). Brain and cognitive evolution: Forms of modularity and functions of mind. *Psychological Bulletin, 128,* 667–698.

Gentile, D. (1998). An ecological approach to the development of perception of emotion in music (Doctoral dissertation, University of Minnesota, Twin Cities Campus, 1998). *Dissertation Abstracts International, 59,* 2454.

*Gérard, C., & Clément, J. (1998). The structure and development of French prosodic representations. *Language and Speech, 41,* 117–142.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston: Houghton-Mifflin.

Giese-Davis, J., & Spiegel, D. (2003). Emotional expression and cancer progression. In R. J. Davidson, H. H. Goldsmith, & K. R. Scherer (Eds.), *Handbook of affective sciences* (pp. 1053–1082). New York: Oxford University Press.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103,* 650– 669.

Giomo, C. J. (1993). An experimental study of children's sensitivity to mood in music. *Psychology of Music, 21,* 141–162.

Gobl, C., & Ní Chasaide, A. (2000). Testing affective correlates of voice quality through analysis and resynthesis. In R. Cowie, E. Douglas-Cowie, & M. Schröder (Eds.), *Proceedings of the ISCA workshop on speech and emotion* [CD-ROM]. Belfast, Ireland: International Speech Communication Association.

Goodall, J. (1986). *The chimpanzees of Gombe: Patterns of behavior.* Cambridge, MA: Harvard University Press.

*Graham, C. R., Hamblin, A. W., & Feldstein, S. (2001). Recognition of emotion in English voices by speakers of Japanese, Spanish and English. *International Review of Applied Linguistics in Language Teaching, 39,* 19–37.

Greasley, P., Sherrard, C., & Waterman, M. (2000). Emotion in language and speech: Methodological issues in naturalistic settings. *Language and Speech, 43,* 355–375.

Gregory, A. H. (1997). The roles of music in society: The ethnomusicological perspective. In D. J. Hargreaves & A. C. North (Eds.), *The social psychology of music* (pp. 123–140). Oxford, England: Oxford University Press.

Gordon, I. E. (1989). *Theories of visual perception.* New York: Wiley.

*Guidetti, M. (1991). L'expression vocale des émotions: Approche interculturelle et développementale [Vocal expression of emotions: A cross-cultural and developmental approach]. *Année Psychologique, 91,* 383– 396.

Gundlach, R. H. (1935). Factors determining the characterization of musical phrases. *American Journal of Psychology, 47,* 624–644.

Hall, J. A., Carter, J. D., & Horgan, T. G. (2001). Gender differences in the nonverbal communication of emotion. In A. Fischer (Ed.), *Gender and emotion* (pp. 97–117). Cambridge, England: Cambridge University Press.

Hammond, K. R., & Stewart, T. R. (Eds.). (2001). *The essential Brunswik: Beginnings, explications, applications.* New York: Oxford University Press.

Handel, S. (1991). *Listening: An introduction to the perception of auditory events.* Cambridge, MA: MIT Press.

Hargreaves, W. A., Starkweather, J. A., & Blacker, K. H. (1965). Voice quality in depression. *Journal of Abnormal Psychology, 70,* 218–220.

Harris, P. L. (1989). *Children and emotion.* Oxford, England: Blackwell.

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1994). *Emotional contagion.* New York: Cambridge University Press.

Hatfield, E., & Rapson, R. L. (2000). Love and attachment processes. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 654–662). New York: Guilford Press.

Hauser, M. D. (2000). The sound and the fury: Primate vocalizations as reflections of emotion and thought. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 77–102). Cambridge, MA: MIT Press.

Havrdova, Z., & Moravek, M. (1979). Changes of the voice expression during suggestively influenced states of experiencing. *Activitas Nervosa Superior, 21,* 33–35.

Helmholtz, H. L. F. von. (1954). *On the sensations of tone as a psycho-logical basis for the theory of music.* New York: Dover. (Original work published 1863)

Hevner, K. (1937). The affective value of pitch and tempo in music. *American Journal of Psychology, 49,* 621–630.

Hietanen, J. K., Surakka, V., & Linnankoski, I. (1998). Facial electromyographic responses to vocal affect expressions. *Psychophysiology, 35,* 530–536.

Höffe, W. L. (1960). Über Beziehungen von Sprachmelodie und Lautstärke [The relationship between speech melody and sound level]. *Phonetica, 5,* 129–159.

*House, D. (1990). On the perception of mood in speech: Implications for the hearing impaired. In L. Eriksson & P. Touati (Eds.), *Working papers* (No. 36, 99–108). Lund, Sweden: Lund University, Department of Linguistics.

Hursch, C. J., Hammond, K. R., & Hursch, J. L. (1964). Some methodological considerations in multiple-cue probability studies. *Psychological Review, 71,* 42–60.

Huttar, G. L. (1968). Relations between prosodic variables and emotions in normal American English utterances. *Journal of Speech and Hearing Research, 11,* 481–487.

Iida, A., Campbell, N., Iga, S., Higuchi, F., & Yasamura, M. (2000). A speech synthesis system with emotion for assisting communication. In R. Cowie, E. Douglas-Cowie, & M. Schröder (Eds.), *Proceedings of the ISCA workshop on speech and emotion* [CD-ROM]. Belfast, Ireland: International Speech Communication Association.

Iriondo, I., Guaus, R., Rodriguez, A., Lazaro, P., Montoya, N., Blanco, J. M., et al. (2000). Validation of an acoustical modelling of emotional expression in Spanish using speech synthesis techniques. In R. Cowie, E. Douglas-Cowie, & M. Schröder (Eds.), *Proceedings of the ISCA workshop on speech and emotion* [CD-ROM]. Belfast, Ireland: International Speech Communication Association.

Izard, C. E. (1992). Basic emotions, relations among emotions, and emotion–cognition relations. *Psychological Review, 99,* 561–565.

Izard, C. E. (1993). Organizational and motivational functions of discrete emotions. In M. Lewis & J. M. Haviland (Eds.), *Handbook of emotions* (pp. 631–641). New York: Guilford Press.

Izard, C. E. (1994). Innate and universal facial expressions: Evidence from developmental and cross-cultural research. *Psychological Bulletin, 115,* 288–299.

Jansens, S., Bloothooft, G., & de Krom, G. (1997). Perception and acoustics of emotions in singing. In *Proceedings of the Fifth European Conference on Speech Communication and Technology: Vol. IV. Eurospeech 97* (pp. 2155–2158). Rhodes, Greece: European Speech Communication Association.

Jo, C.-W., Ferencz, A., & Kim, D.-H. (1999). Experiments regarding the superposition of emotional features on neutral Korean speech. *Lecture Notes in Artificial Intelligence, 1692,* 333–336.

*Johnson, W. F., Emde, R. N., Scherer, K. R., & Klinnert, M. D. (1986). Recognition of emotion from vocal cues. *Archives of General Psychiatry, 43,* 280–283.

Johnson-Laird, P. N., & Oatley, K. (1989). The language of emotions: An analysis of a semantic field. *Cognition & Emotion, 3,* 81–123.

Johnson-Laird, P. N., & Oatley, K. (1992). Basic emotions, rationality, and folk theory. *Cognition & Emotion, 6,* 201–223.

Johnstone, T. (2001). *The communication of affect through modulation of non-verbal vocal parameters.* Unpublished doctoral dissertation, University of Western Australia, Nedlands, Western Australia.

Johnstone, T., & Scherer, K. R. (1999). The effects of emotions on voice quality. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A. C. Bailey (Eds.), *Proceedings of the XIVth International Congress of Phonetic Sciences* [CD-ROM]. Berkeley: University of California, Department of Linguistics.

Johnstone, T., & Scherer, K. R. (2000). Vocal communication of emotion.

In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 220–235). New York: Guilford Press.

Jürgens, U. (1976). Projections from the cortical larynx area in the squirrel monkey. *Experimental Brain Research, 25,* 401–411.

Jürgens, U. (1979). Vocalization as an emotional indicator: A neuroethological study in the squirrel monkey. *Behaviour, 69,* 88–117.

Jürgens, U. (1992). On the neurobiology of vocal communication. In H. Papoušek, U. Jürgens, & M. Papoušek (Eds.), *Nonverbal vocal communication: Comparative and developmental approaches* (pp. 31–42). Cambridge, England: Cambridge University Press.

Jürgens, U. (2002). Neural pathways underlying vocal control. *Neuroscience and Biobehavioral Reviews, 26,* 235–258.

Jürgens, U., & von Cramon, D. (1982). On the role of the anterior cingulate cortex in phonation: A case report. *Brain and Language, 15,* 234–248.

Juslin, P. N. (1993). *The influence of expressive intention on electric guitar performance.* Unpublished bachelor's thesis, Uppsala University, Uppsala, Sweden.

Juslin, P. N. (1995). Emotional communication in music viewed through a Brunswikian lens. In G. Kleinen (Ed.), *Musical expression. Proceedings of the Conference of ESCOM and DGM 1995* (pp. 21–25). University of Bremen: Bremen, Germany.

Juslin, P. N. (1996). Affective computing. *Ung Forskning, 4,* 60–64.

*Juslin, P. N. (1997a). Can results from studies of perceived expression in musical performance be generalized across response formats? *Psychomusicology, 16,* 77–101.

*Juslin, P. N. (1997b). Emotional communication in music performance: A functionalist perspective and some data. *Music Perception, 14,* 383–418.

*Juslin, P. N. (1997c). Perceived emotional expression in synthesized performances of a short melody: Capturing the listener's judgment policy. *Musicae Scientiae, 1,* 225–256.

Juslin, P. N. (1998). A functionalist perspective on emotional communication in music performance (Doctoral dissertation, Uppsala University, 1998). In *Comprehensive summaries of Uppsala dissertations from the faculty of social sciences, No. 78* (pp. 7–65). Uppsala, Sweden: Uppsala University Library.

Juslin, P. N. (2000). Cue utilization in communication of emotion in music performance: Relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1797–1813.

Juslin, P. N. (2001a). A Brunswikian approach to emotional communication in music performance. In K. R. Hammond & T. R. Stewart (Eds.), *The essential Brunswik: Beginnings, explications, applications* (pp. 426–430). New York: Oxford University Press.

Juslin, P. N. (2001b). Communicating emotion in music performance. A review and a theoretical framework. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 309–337). New York: Oxford University Press.

Juslin, P. N. (2003). Five facets of musical expression: A psychologist's perspective on music performance. *Psychology of Music, 31,* 273–302.

Juslin, P. N. (in press). Vocal expression and musical expression: Parallels and contrasts. In A. Kappas (Ed.), *Proceedings of the 11th Meeting of the International Society for Research on Emotions.* Quebec City, Quebec, Canada: International Society for Research on Emotions.

Juslin, P. N., Friberg, A., & Bresin, R. (2002). Toward a computational model of expression in music performance: The GERM model. *Musicae Scientiae, Special Issue 2001–2002,* 63–122.

Juslin, P. N., Friberg, A., Schoonderwaldt, E., & Karlsson, J. (in press). Feedback-learning of musical expressivity. In A. Williamon (Ed.), *Enhancing musical performance: A resource for performers, teachers, and researchers.* New York: Oxford University Press.

*Juslin, P. N., & Laukka, P. (2000). Improving emotional communication in music performance through cognitive feedback. *Musicae Scientiae, 4,* 151–183.

*Juslin, P. N., & Laukka, P. (2001). Impact of intended emotion intensity on decoding accuracy and cue utilization in vocal expression of emotion. *Emotion, 1,* 381–412.

Juslin, P. N., & Laukka, P. (in press). Emotional expression in speech and music: Evidence of cross-modal similarities. *Annals of the New York Academy of Sciences.* New York: New York Academy of Sciences.

Juslin, P. N., & Lindström, E. (2003). *Musical expression of emotions: Modeling composed and performed features.* Unpublished manuscript.

*Juslin, P. N., & Madison, G. (1999). The role of timing patterns in recognition of emotional expression from musical performance. *Music Perception, 17,* 197–221.

Juslin, P. N., & Persson, R. S. (2002). Emotional communication. In R. Parncutt & G. E. McPherson (Eds.), *The science and psychology of music performance: Creative strategies for teaching and learning* (pp. 219–236). New York: Oxford University Press.

Juslin, P. N., & Sloboda, J. A. (Eds.). (2001). *Music and emotion: Theory and research.* New York: Oxford University Press.

Juslin, P. N., & Zentner, M. R. (2002). Current trends in the study of music and emotion: Overture. *Musicae Scientiae, Special Issue 2001–2002,* 3–21.

Kaiser, L. (1962). Communication of affects by single vowels. *Synthese, 14,* 300–319.

Kaiser, S., & Scherer, K. R. (1998). Models of "normal" emotions applied to facial and vocal expression in clinical disorders. In W. F. Flack & J. D. Laird (Eds.), *Emotions in psychopathology: Theory and research* (pp. 81–98). New York: Oxford University Press.

Kastner, M. P., & Crowder, R. G. (1990). Perception of the major/minor distinction: IV. Emotional connotations in young children. *Music Perception, 8,* 189–202.

Katz, G. S. (1997). *Emotional speech: A quantitative study of vocal acoustics in emotional expression.* Unpublished doctoral dissertation, University of Pittsburgh, Pittsburgh, PA.

Katz, G. S., Cohn, J. F., & Moore, C. A. (1996). A combination of vocal F_0 dynamic and summary features discriminates between three pragmatic categories of infant-directed speech. *Child Development, 67,* 205–217.

Keltner, D., & Gross, J. J. (1999). Functional accounts of emotions. *Cognition & Emotion, 13,* 465–466.

Kienast, M., & Sendlmeier, W. F. (2000). Acoustical analysis of spectral and temporal changes in emotional speech. In R. Cowie, E. Douglas-Cowie, & M. Schröder (Eds.), *Proceedings of the ISCA workshop on speech and emotion* [CD-ROM]. Belfast, Ireland: International Speech Communication Association.

*Kitahara, Y., & Tohkura, Y. (1992). Prosodic control to express emotions for man-machine speech interaction. *IEICE Transactions on Fundamentals of Electronics, Communications, and Computer Sciences, 75,* 155–163.

Kivy, P. (1980). *The corded shell.* Princeton, NJ: Princeton University Press.

Klasmeyer, G., & Sendlmeier, W. F. (1997). The classification of different phonation types in emotional and neutral speech. *Forensic Linguistics, 1,* 104–124.

*Knower, F. H. (1941). Analysis of some experimental variations of simulated vocal expressions of the emotions. *Journal of Social Psychology, 14,* 369–372.

*Knower, F. H. (1945). Studies in the symbolism of voice and action: V. The use of behavioral and tonal symbols as tests of speaking achievement. *Journal of Applied Psychology, 29,* 229–235.

*Konishi, T., Imaizumi, S., & Niimi, S. (2000). Vibrato and emotion in singing voice. In C. Woods, G. Luck, R. Brochard, F. Seddon, & J. A. Sloboda (Eds.), *Proceedings of the Sixth International Conference for Music Perception and Cognition* [CD-ROM]. Keele, England: Keele University.

*Kotlyar, G. M., & Morozov, V. P. (1976). Acoustical correlates of the

emotional content of vocalized speech. *Soviet Physics: Acoustics, 22,* 208–211.

*Kramer, E. (1964). Elimination of verbal cues in judgments of emotion from voice. *Journal of Abnormal and Social Psychology, 68,* 390–396.

Kratus, J. (1993). A developmental study of children's interpretation of emotion in music. *Psychology of Music, 21,* 3–19.

Krebs, J. R., & Dawkins, R. (1984). Animal signals: Mind-reading and manipulation. In J. R. Krebs & N. B. Davies, (Eds.), *Behavioural ecology: An evolutionary approach* (2nd ed., pp. 380–402). Oxford, England: Blackwell.

Kuny, S., & Stassen, H. H. (1993). Speaking behavior and voice sound characteristics in depressive patients during recovery. *Journal of Psychiatric Research, 27,* 289–307.

Kuroda, I., Fujiwara, O., Okamura, N., & Utsuki, N. (1976). Method for determining pilot stress through analysis of voice communication. *Aviation, Space, and Environmental Medicine, 47,* 528–533.

Kuypers, H. G. (1958). Corticobulbar connections to the pons and lower brain stem in man. *Brain, 81,* 364–388.

Ladd, D. R., Silverman, K. E. A., Tolkmitt, F., Bergmann, G., & Scherer, K. R. (1985). Evidence of independent function of intonation contour type, voice quality, and F0 range in signaling speaker affect. *Journal of the Acoustical Society of America, 78,* 435–444.

Langeheinecke, E. J., Schnitzler, H.-U., Hischer-Buhrmester, U., & Behne, K.-E. (1999, March). *Emotions in the singing voice: Acoustic cues for joy, fear, anger and sadness.* Poster session presented at the Joint Meeting of the Acoustical Society of America and the Acoustical Society of Germany, Berlin.

Langer, S. (1951). *Philosophy in a new key* (2nd ed.). New York: New American Library.

Laukka, P. (in press). Categorical perception of emotions in vocal expression. *Annals of the New York Academy of Sciences.* New York: New York Academy of Sciences.

*Laukka, P., & Gabrielsson, A. (2000). Emotional expression in drumming performance. *Psychology of Music, 28,* 181–189.

Laukka, P., Juslin, P. N., & Bresin, R. (2003). *A dimensional approach to vocal expression of emotion.* Manuscript submitted for publication.

Laukkanen, A.-M., Vilkman, E., Alku, P., & Oksanen, H. (1996). Physical variations related to stress and emotional state: A preliminary study. *Journal of Phonetics, 24,* 313–335.

Laukkanen, A.-M., Vilkman, E., Alku, P., & Oksanen, H. (1997). On the perception of emotions in speech: The role of voice quality. *Logopedics Phoniatrics Vocology, 22,* 157–168.

Laver, J. (1980). *The phonetic description of voice quality.* Cambridge, England: Cambridge University Press.

Lazarus, R. S. (1991). *Emotion and adaptation.* New York: Oxford University Press.

Lee, I. J. (1939). Some conceptions of emotional appeal in rhetorical theory. *Speech Monographs, 6,* 66–86.

*Leinonen, L., Hiltunen, T., Linnankoski, I., & Laakso, M.-L. (1997). Expression of emotional-motivational connotations with a one-word utterance. *Journal of the Acoustical Society of America, 102,* 1853–1863.

*Léon, P. R. (1976). De l'analyse psychologique a la catégorisation auditive et acoustique des émotions dans la parole [On the psychological analysis of auditory and acoustic categorization of emotions in speech]. *Journal de Psychologie Normale et Pathologique, 73,* 305–324.

Levenson, R. W. (1992). Autonomic nervous system differences among emotions. *Psychological Science, 3,* 23–27.

Levenson, R. W. (1994). Human emotion: A functional view. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 123–126). New York: Oxford University Press.

Levin, H., & Lord, W. (1975). Speech pitch frequency as an emotional state indicator. *IEEE Transactions on Systems, Man, and Cybernetics, 5,* 259–273.

*Levitt, E. A. (1964). The relationship between abilities to express emotional meanings vocally and facially. In J. R. Davitz (Ed.), *The communication of emotional meaning* (pp. 87–100). New York: McGraw-Hill.

Levman, B. G. (2000). Western theories of music origin, historical and modern. *Musicae Scientiae, 4,* 185–211.

Lewis, M. (2000). The emergence of human emotions. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 265–280). New York: Guilford Press.

Lieberman, P. (1961). Perturbations in vocal pitch. *Journal of the Acoustical Society of America, 33,* 597–603.

Lieberman, P., & Michaels, S. B. (1962). Some aspects of fundamental frequency and envelope amplitude as related to the emotional content of speech. *Journal of the Acoustical Society of America, 34,* 922–927.

Lindström, E., Juslin, P. N., Bresin, R., & Williamon, A. (2003). Expressivity comes from within your soul: A questionnaire study of music students' perspectives on expressivity. *Research Studies in Music Education, 20,* 23–47.

London, J. (2002). Some theories of emotion in music and their implications for research in music psychology. *Musicae Scientiae, Special Issue 2001–2002,* 23–36.

Lundqvist, L. G., Carlsson, F., & Hilmersson, P. (2000, July). *Facial electromyography, autonomic activity, and emotional experience to happy and sad music.* Paper presented at the 27th International Congress of Psychology, Stockholm, Sweden.

MacLean, P. (1993). Cerebral evolution of emotion. In M. Lewis & J. M. Haviland (Eds.), *Handbook of emotions* (pp. 67–83). New York: Guilford Press.

Madison, G. (2000a). Interaction between melodic structure and performance variability on the expressive dimensions perceived by listeners. In C. Woods, G. Luck, R. Brochard, F. Seddon, & J. A. Sloboda (Eds.), *Proceedings of the Sixth International Conference on Music Perception and Cognition, August 2000* [CD-ROM]. Keele, England: Keele University.

Madison, G. (2000b). Properties of expressive variability patterns in music performances. *Journal of New Music Research, 29,* 335–356.

Markel, N. N., Bein, M. F., & Phillis, J. A. (1973). The relationship between words and tone-of-voice. *Language and Speech, 16,* 15–21.

Marler, P. (1977). The evolution of communication. In T. A. Sebeok (Ed.), *How animals communicate* (pp. 45–70). Bloomington: Indiana University Press.

Mastropieri, D., & Turkewitz, G. (1999). Prenatal experience and neonatal responsiveness to vocal expressions of emotion. *Developmental Psychobiology, 35,* 204–214.

Mayne, T. J. (2001). Emotions and health. In T. J. Mayne & G. A. Bonanno (Eds.), *Emotions: Current issues and future directions* (pp. 361–397). New York: Guilford Press.

McCluskey, K. W., & Albas, D. C. (1981). Perception of the emotional content of speech by Canadian and Mexican children, adolescents, and adults. *International Journal of Psychology, 16,* 119–132.

McCluskey, K. W., Albas, D. C., Niemi, R. R., Cuevas, C., & Ferrer, C. A. (1975). Cross-cultural differences in the perception of emotional content of speech: A study of the development of sensitivity in Canadian and Mexican children. *Developmental Psychology, 11,* 551–555.

McRoberts, G. W., Studdert-Kennedy, M., & Shankweiler, D. P. (1995). The role of fundamental frequency in signaling linguistic stress and affect: Evidence for a dissociation. *Perception & Psychophysics, 57,* 159–174.

*Mergl, R., Piesbergen, C., & Tunner, W. (1998). Musikalisch-improvisatorischer Ausdruck und erkennen von gefühlsqualitäten [Expression in musical improvisation and the recognition of emotional qualities]. *Musikpsychologie: Jahrbuch der Deutschen Gesellschaft für Musikpsychologie, 13,* 69–81.

Metfessel, M. (1932). The vibrato in artistic voices. In C. E. Seashore

(Ed.), *University of Iowa studies in the psychology of music: Vol. 1. The vibrato* (pp. 14–117). Iowa City: University of Iowa Press.

Meyer, L. B. (1956). *Emotion and meaning in music.* Chicago: Chicago University Press.

Moriyama, T., & Ozawa, S. (2001). Measurement of human vocal emotion using fuzzy control. *Systems and Computers in Japan, 32,* 59–68.

Morozov, V. P. (1996). Emotional expressiveness of the singing voice: The role of macrostructural and microstructural modifications of spectra. *Logopedics Phoniatrics Vocology, 21,* 49–58.

Morton, E. S. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *American Naturalist, 111,* 855–869.

Morton, J. B., & Trehub, S. E. (2001). Children's understanding of emotion in speech. *Child Development, 72,* 834–843.

*Mozziconacci, S. J. L. (1998). *Speech variability and emotion: Production and perception.* Eindhoven, the Netherlands: Technische Universiteit Eindhoven.

Murray, I. R., & Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *Journal of the Acoustical Society of America, 93,* 1097–1108.

Murray, I. R., & Arnott, J. L. (1995). Implementation and testing of a system for producing emotion-by-rule in synthetic speech. *Speech Communication, 16,* 369–390.

Murray, I. R., Baber, C., & South, A. (1996). Towards a definition and working model of stress and its effects on speech. *Speech Communication, 20,* 3–12.

Nair, D. G., Large, E. W., Steinberg, F., & Kelso, J. A. S. (2002). Perceiving emotion in expressive piano performance: A functional MRI study. In K. Stevens, D. Burnham, G. McPherson, E. Schubert, & J. Renwick (Eds.), *Proceedings of the 7th International Conference on Music Perception and Cognition, July 2002* [CD-ROM]. Adelaide, South Australia: Causal Productions.

Nawrot, E. S. (2003). The perception of emotional expression in music: Evidence from infants, children, and adults. *Psychology of Music, 31,* 75–92.

Neumann, R., & Strack, F. (2000). Mood contagion: The automatic transfer of mood between persons. *Journal of Personality and Social Psychology, 79,* 211–223.

Niedenthal, P. M., & Showers, C. (1991). The perception and processing of affective information and its influences on social judgment. In J. P. Forgas (Ed.), *Emotion & social judgments* (pp. 125–143). Oxford, England: Pergamon Press.

Nilsonne, Å. (1987). Acoustic analysis of speech variables during depression and after improvement. *Acta Psychiatrica Scandinavica, 76,* 235–245.

Novak, A., & Vokral, J. (1993). Emotions in the sight of long-time averaged spectrum and three-dimensional analysis of periodicity. *Folia Phoniatrica, 45,* 198–203.

Oatley, K. (1992). *Best laid schemes.* Cambridge, MA: Harvard University Press.

Oatley, K., & Jenkins, J. M. (1996). *Understanding emotions.* Oxford, England: Blackwell.

Ohala, J. J. (1983). Cross-language use of pitch: An ethological view. *Phonetica, 40,* 1–18.

Ohgushi, K., & Hattori, M. (1996a, December). *Acoustic correlates of the emotional expression in vocal performance.* Paper presented at the third joint meeting of the Acoustical Society of America and Acoustical Society of Japan, Honolulu, HI.

Ohgushi, K., & Hattori, M. (1996b). Emotional communication in performance of vocal music. In B. Pennycook & E. Costa-Giomi (Eds.), *Proceedings of the Fourth International Conference on Music Perception and Cognition* (pp. 269–274). Montreal, Quebec, Canada: McGill University.

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness:

Toward an evolved module of fear and fear learning. *Psychological Review, 108,* 483–522.

Ortony, A., & Turner, T. J. (1990). What's basic about basic emotions? *Psychological Review, 97,* 315–331.

*Oura, Y., & Nakanishi, R. (2000). How do children and college students recognize emotions of piano performances? *Journal of Music Perception and Cognition, 6,* 13–29.

Owren, M. J., & Bachorowski, J.-A. (2001). The evolution of emotional expression: A "selfish-gene" account of smiling and laughter in early hominids and humans. In T. J. Mayne & G. A. Bonanno (Eds.), *Emotions: Current issues and future directions* (pp. 152–191). New York: Guilford Press.

Paeschke, A., & Sendlmeier, W. F. (2000). Prosodic characteristics of emotional speech: Measurements of fundamental frequency. In R. Cowie, E. Douglas-Cowie, & M. Schröder (Eds.), *Proceedings of the ISCA workshop on speech and emotion* [CD-ROM]. Belfast, Ireland: International Speech Communication Association.

Palmer, C. (1997). Music performance. *Annual Review of Psychology, 48,* 115–138.

Panksepp, J. (1985). Mood changes. In P. J. Vinken, G. W. Bruyn, & H. L. Klawans (Eds.), *Handbook of clinical neurology: Vol. 1. Clinical neuropsychology.* (pp. 271–285). Amsterdam: Elsevier.

Panksepp, J. (1992). A critical role for affective neuroscience in resolving what is basic about basic emotions. *Psychological Review, 99,* 554–560.

Panksepp, J. (1998). *Affective neuroscience.* New York: Oxford University Press.

Panksepp, J. (2000). Emotions as natural kinds within the mammalian brain. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 137–156). New York: Guilford Press.

Panksepp, J., & Panksepp, J. B. (2000). The seven sins of evolutionary psychology. *Evolution and Cognition, 6,* 108–131.

Papoušek, H., Jürgens, U., & Papoušek, M. (Eds.). (1992). *Nonverbal vocal communication: Comparative and developmental approaches.* Cambridge, England: Cambridge University Press.

Papoušek, M. (1996). Intuitive parenting: A hidden source of musical stimulation in infancy. In I. Deliége & J. A. Sloboda (Eds.), *Musical beginnings: Origins and development of musical competence* (pp. 89–112). Oxford, England: Oxford University Press.

Patel, A. D., & Peretz, I. (1997). Is music autonomous from language? A neuropsychological appraisal. In I. Deliége & J. A. Sloboda (Eds.), *Perception and cognition of music* (pp. 191–215). Hove, England: Psychology Press.

Patel, A. D., Peretz, I., Tramo, M., & Labreque, R. (1998). Processing prosodic and musical patterns: A neuropsychological investigation. *Brain and Language, 61,* 123–144.

Pell, M. D. (2001). Influence of emotion and focus location on prosody in matched statements and questions. *Journal of the Acoustical Society of America, 109,* 1668–1680.

Peretz, I. (2001). Listen to the brain: A biological perspective on musical emotions. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 105–134). New York: Oxford University Press.

Peretz, I. (2002). Brain specialization for music. *Neuroscientist, 8,* 372–380.

Peretz, I., Gagnon, L., & Bouchard, B. (1998). Music and emotion: Perceptual determinants, immediacy, and isolation after brain damage. *Cognition, 68,* 111–141.

*Pfaff, P. L. (1954). An experimental study of the communication of feeling without contextual material. *Speech Monographs, 21,* 155–156.

Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: A meta analysis of emotion activation studies in PET and fMRI. *NeuroImage, 16,* 331–348.

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

Planalp, S. (1998). Communicating emotion in everyday life: Cues, chan-

nels, and processes. In P. A. Andersen & L. K. Guerrero (Eds.), *Handbook of communication and emotion* (pp. 29–48). New York: Academic Press.

Ploog, D. (1981). Neurobiology of primate audio-vocal behavior. *Brain Research Reviews, 3,* 35–61.

Ploog, D. (1986). Biological foundations of the vocal expressions of emotions. In R. Plutchik & H. Kellerman (Eds.), *Emotion: Theory, research, and experience: Vol. 3. Biological foundations of emotion* (pp. 173–197). New York: Academic Press.

Ploog, D. (1992). The evolution of vocal communication. In H. Papoušek, U. Jürgens, & M. Papoušek (Eds.), *Nonverbal vocal communication: Comparative and developmental approaches* (pp. 6–30). Cambridge, England: Cambridge University Press.

Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In R. Plutchik & H. Kellerman (Eds.), *Emotion: Theory, research, and experience: Vol. 1. Theories of emotion* (pp. 3–33). New York: Academic Press.

Plutchik, R. (1994). *The psychology and biology of emotion.* New York: Harper-Collins.

Pollack, I., Rubenstein, H., & Horowitz, A. (1960). Communication of verbal modes of expression. *Language and Speech, 3,* 121–130.

Power, M., & Dalgleish, T. (1997). *Cognition and emotion: From order to disorder.* Hove, England: Psychology Press.

Protopapas, A., & Lieberman, P. (1997). Fundamental frequency of phonation and perceived emotional stress. *Journal of the Acoustical Society of America, 101,* 2267–2277.

Rapoport, E. (1996). Emotional expression code in opera and lied singing. *Journal of New Music Research, 25,* 109–149.

Richman, B. (1987). Rhythm and melody in gelada vocal exchanges. *Primates, 28,* 199–223.

Rigg, M. G. (1940). The effect of register and tonality upon musical mood. *Journal of Musicology, 2,* 49–61.

Roessler, R., & Lester, J. W. (1976). Voice predicts affect during psychotherapy. *Journal of Nervous and Mental Disease, 163,* 166–176.

Rosenthal, R. (1982). Judgment studies. In K. R. Scherer & P. Ekman (Eds.), *Handbook of methods in nonverbal behavior research* (pp. 287–361). Cambridge, England: Cambridge University Press.

Rosenthal, R., & Rubin, D. B. (1989). Effect size estimation for one-sample multiple-choice-type data: Design, analysis, and meta-analysis. *Psychological Bulletin, 106,* 332–337.

Ross, B. H., & Spalding, T. L. (1994). Concepts and categories. In R. J. Sternberg (Ed.), *Thinking and problem solving* (2nd ed., pp. 119–150). New York: Academic Press.

Rousseau, J. J. (1986). Essay on the origin of languages. In J. H. Moran & A. Gode (Eds.), *On the origin of language: Two essays* (pp. 5–74). Chicago: University of Chicago Press. (Original work published 1761)

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39,* 1161–1178.

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin, 115,* 102–141.

Russell, J. A., Bachorowski, J.-A., & Fernández-Dols, J.-M. (2003). Facial and vocal expressions of emotion. *Annual Review of Psychology, 54,* 329–349.

Salgado, A. G. (2000). Voice, emotion and facial gesture in singing. In C. Woods, G. Luck, R. Brochard, F. Seddon, & J. A. Sloboda (Eds.), *Proceedings of the Sixth International Conference for Music Perception and Cognition* [CD-ROM]. Keele, England: Keele University.

Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition, and Personality, 9,* 185–211.

Saperston, B., West, R., & Wigram, T. (1995). *The art and science of music therapy: A handbook.* Chur, Switzerland: Harwood.

Scherer, K. R. (1974). Acoustic concomitants of emotional dimensions: Judging affect from synthesized tone sequences. In S. Weitz (Ed.), *Nonverbal communication* (pp. 105–111). New York: Oxford University Press.

Scherer, K. R. (1978). Personality inference from voice quality: The loud voice of extroversion. *European Journal of Social Psychology, 8,* 467–487.

Scherer, K. R. (1979). Non-linguistic vocal indicators of emotion and psychopathology. In C. E. Izard (Ed.), *Emotions in personality and psychopathology* (pp. 493–529). New York: Plenum Press.

Scherer, K. R. (1982). Methods of research on vocal communication: Paradigms and parameters. In K. R. Scherer & P. Ekman (Eds.), *Handbook of methods in nonverbal behavior research* (pp. 136–198). Cambridge, England: Cambridge University Press.

Scherer, K. R. (1985). Vocal affect signalling: A comparative approach. In J. Rosenblatt, C. Beer, M.-C. Busnel, & P. J. B. Slater (Eds.), *Advances in the study of behavior* (Vol. 15, pp. 189–244). New York: Academic Press.

Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin, 99,* 143–165.

Scherer, K. R. (1989). Vocal correlates of emotional arousal and affective disturbance. In H. Wagner & A. Manstead (Eds.), *Handbook of social psychophysiology* (pp. 165–197). New York: Wiley.

Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice, 9,* 235–248.

Scherer, K. R. (2000). Psychological models of emotion. In J. Borod (Ed.), *The neuropsychology of emotion* (pp. 137–162). New York: Oxford University Press.

Scherer, K. R. (2001). Appraisal considered as a process of multi-level sequential checking. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research* (pp. 92–120). New York: Oxford University Press.

*Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology, 32,* 76–92.

*Scherer, K. R., Banse, R., Wallbott, H. G., & Goldbeck, T. (1991). Vocal cues in emotion encoding and decoding. *Motivation and Emotion, 15,* 123–148.

Scherer, K. R., & Oshinsky, J. S. (1977). Cue utilisation in emotion attribution from auditory stimuli. *Motivation and Emotion, 1,* 331–346.

Scherer, K. R., & Zentner, M. R. (2001). Emotional effects of music: Production rules. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 361–392). New York: Oxford University Press.

*Schröder, M. (1999). Can emotions be synthesized without controlling voice quality? *Phonus, 4,* 35–50.

*Schröder, M. (2000). Experimental study of affect bursts. In R. Cowie, E. Douglas-Cowie, & M. Schröder (Eds.), *Proceedings of the ISCA workshop on speech and emotion* [CD-ROM]. Belfast, Ireland: International Speech Communication Association.

Schröder, M. (2001). Emotional speech synthesis: A review. In *Proceedings of the 7th European Conference on Speech Communication and Technology: Vol. 1. Eurospeech 2001, September 3–7, 2001* (pp. 561–564). Aalborg, Denmark: International Speech Communication Association.

Schwartz, G. E., Weinberger, D. A., & Singer, J. A. (1981). Cardiovascular differentiation of happiness, sadness, anger, and fear following imagery and exercise. *Psychosomatic Medicine, 43,* 343–364.

Scott, J. P. (1980). The function of emotions in behavioral systems: A systems theory analysis. In R. Plutchik & H. Kellerman (Eds.), *Emotion: Theory, research, and experience: Vol. 1. Theories of emotion* (pp. 35–56). New York: Academic Press.

Seashore, C. E. (1927). Phonophotography in the measurement of the expression of emotion in music and speech. *Scientific Monthly, 24,* 463–471.

Seashore, C. E. (1947). *In search of beauty in music: A scientific approach to musical aesthetics.* Westport, CT: Greenwood Press.

Sedláček, K., & Sychra, A. (1963). Die Melodie als Faktor des emotionellen Ausdrucks [Melody as a factor in emotional expression]. *Folia Phoniatrica, 15,* 89–98.

Senju, M., & Ohgushi, K. (1987). How are the player's ideas conveyed to the audience? *Music Perception, 4,* 311–324.

Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication.* Urbana: University of Illinois Press.

Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion knowledge: Further explorations of a prototype approach. *Journal of Personality and Social Psychology, 52,* 1061–1086.

Sherman, M. (1928). Emotional character of the singing voice. *Journal of Experimental Psychology, 11,* 495–497.

Shields, S. A. (1984). Distinguishing between emotion and non-emotion: Judgments about experience. *Motivation and Emotion, 8,* 355–369.

Siegman, A. W., Anderson, R. A., & Berger, T. (1990). The angry voice: Its effects on the experience of anger and cardiovascular reactivity. *Psychosomatic Medicine, 52,* 631–643.

Siegwart, H., & Scherer, K. R. (1995). Acoustic concomitants of emotional expression in operatic singing: The case of Lucia in Ardi gli incensi. *Journal of Voice, 9,* 249–260.

Simonov, P. V., Frolov, M. V., & Taubkin, V. L. (1975). Use of the invariant method of speech analysis to discern the emotional state of announcers. *Aviation, Space, and Environmental Medicine, 46,* 1014–1016.

Singh, L., Morgan, J. L., & Best, C. T. (2002). Infants' listening preferences: Baby talk or happy talk? *Infancy, 3,* 365–394.

Skinner, E. R. (1935). A calibrated recording and analysis of the pitch, force and quality of vocal tones expressing happiness and sadness. And a determination of the pitch and force of the subjective concepts of ordinary, soft and loud tones. *Speech Monographs, 2,* 81–137.

Sloboda, J. A., & Juslin, P. N. (2001). Psychological perspectives on music and emotion. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 71–104). New York: Oxford University Press.

Snowdon, C. T. (2003). Expression of emotion in nonhuman animals. In R. J. Davidson, H. H. Goldsmith, & K. R. Scherer (Eds.), *Handbook of affective sciences* (pp. 457–480). New York: Oxford University Press.

Sobin, C., & Alpert, M. (1999). Emotion in speech: The acoustic attributes of fear, anger, sadness, and joy. *Journal of Psycholinguistic Research, 28,* 347–365.

*Sogon, S. (1975). A study of the personality factor which affects the judgement of vocally expressed emotions. *Japanese Journal of Psychology, 46,* 247–254.

Soken, N. H., & Pick, A. D. (1999). Infants' perception of dynamic affective expressions: Do infants distinguish specific expressions? *Child Development, 70,* 1275–1282.

Spencer, H. (1857). The origin and function of music. *Fraser's Magazine, 56,* 396–408.

Stassen, H. H., Kuny, S., & Hell, D. (1998). The speech analysis approach to determining onset of improvement under antidepressants. *European Neuropsychopharmacology, 8,* 303–310.

Steffen-Batóg, M., Madelska, L., & Katulska, K. (1993). The role of voice timbre, duration, speech melody and dynamics in the perception of the emotional colouring of utterances. *Studia Phonetica Posnaniensia, 4,* 73–92.

Stibbard, R. M. (2001). *Vocal expression of emotions in non-laboratory speech: An investigation of the Reading/Leeds Emotion in Speech Project annotation data.* Unpublished doctoral dissertation, University of Reading, Reading, England.

Storr, A. (1992). *Music and the mind.* London: Harper-Collins.

Sulc, J. (1977). To the problem of emotional changes in human voice. *Activitas Nervosa Superior, 19,* 215–216.

Sundberg, J. (1982). Speech, song, and emotions. In M. Clynes (Ed.), *Music, mind, and brain. The neuropsychology of music* (pp. 137–149). New York: Plenum Press.

Sundberg, J. (1991). *The science of musical sounds.* New York: Academic Press.

Sundberg, J. (1999). The perception of singing. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 171–214). San Diego, CA: Academic Press.

Sundberg, J., Iwarsson, J., & Hagegård, H. (1995). A singer's expression of emotions in sung performance. In O. Fujimura & M. Hirano (Eds.), *Vocal fold physiology: Voice quality control* (pp. 217–229). San Diego, CA: Singular Press.

Svejda, M. J. (1982). The development of infant sensitivity to affective messages in the mother's voice (Doctoral dissertation, University of Denver, 1982). *Dissertation Abstracts International, 42,* 4623.

Tartter, V. C. (1980). Happy talk: Perceptual and acoustic effects of smiling on speech. *Perception & Psychophysics, 27,* 24–27.

Tartter, V. C., & Braun, D. (1994). Hearing smiles and frowns in normal and whisper registers. *Journal of the Acoustical Society of America, 96,* 2101–2107.

Terwogt, M. M., & van Grinsven, F. (1988). Recognition of emotions in music by children and adults. *Perceptual and Motor Skills, 67,* 697–698.

Terwogt, M. M., & van Grinsven, F. (1991). Musical expression of mood states. *Psychology of Music, 19,* 99–109.

*Tickle, A. (2000). English and Japanese speakers' emotion vocalisation and recognition: A comparison highlighting vowel quality. In R. Cowie, E. Douglas-Cowie, & M. Schröder (Eds.), *Proceedings of the ISCA workshop on speech and emotion* [CD-ROM]. Belfast, Ireland: International Speech Communication Association.

Tischer, B. (1993). Äusserungsinterne Änderungen des emotionalen Eindrucks mündlicher Sprache: Dimensionen und akustische Korrelate der Eindruckswirkung [Within-utterance variations in the emotional impression of speech: Dimensions and acoustic correlates of perceived emotion]. *Zeitschrift für Experimentelle und Angewandte Psychologie, 40,* 644–675.

Tischer, B. (1995). Acoustic correlates of perceived emotional stress. In I. Trancoso & R. Moore (Eds.), *Proceedings of the ESCA-NATO Tutorial and Research Workshop on Speech Under Stress* (pp. 29–32). Lisbon, Portugal: European Speech Communication Association.

Tolkmitt, F. J., & Scherer, K. R. (1986). Effect of experimentally induced stress on vocal parameters. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 302–313.

Tomkins, S. (1962). *Affect, imagery, and consciousness: Vol. 1. The positive affects.* New York: Springer.

*Trainor, L. J., Austin, C. M., & Desjardins, R. N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychological Science, 11,* 188–195.

van Bezooijen, R. (1984). *Characteristics and recognizability of vocal expressions of emotion.* Dordrecht, the Netherlands: Foris.

*van Bezooijen, R., Otto, S. A., & Heenan, T. A. (1983). Recognition of vocal expressions of emotion: A three-nation study to identify universal characteristics. *Journal of Cross-Cultural Psychology, 14,* 387–406.

Västfjäll, D. (2002). A review of the musical mood induction procedure. *Musicae Scientiae, Special Issue 2001–2002,* 173–211.

Verny, T., & Kelly J. (1981). *The secret life of the unborn child.* New York: Delta.

Von Bismarck, G. (1974). Sharpness as an attribute of the timbre of steady state sounds. *Acustica, 30,* 146–159.

Wagner, H. L. (1993). On measuring performance in category judgment studies of nonverbal behavior. *Journal of Nonverbal Behavior, 17,* 3–28.

*Wallbott, H. G., & Scherer, K. R. (1986). Cues and channels in emotion recognition. *Journal of Personality and Social Psychology, 51,* 690–699.

Wallin, N. L., Merker, B., & Brown, S. (Eds.). (2000). *The origins of music.* Cambridge, MA: MIT Press.

Watson, D. (1991). *The Wordsworth dictionary of musical quotations.* Ware, England: Wordsworth.

Watson, K. B. (1942). The nature and measurement of musical meanings. *Psychological Monographs, 54,* 1–43.

Wedin, L. (1972). Multidimensional study of perceptual-emotional qualities in music. *Scandinavian Journal of Psychology, 13,* 241–257.

Wiley, R. H., & Richards, D. G. (1978). Physical constraints on acoustic communication in the atmosphere: Implications for the evolution of animal vocalizations. *Behavioral Ecology and Sociobiology, 3,* 69–94.

Whiteside, S. P. (1999a). Acoustic characteristics of vocal emotions simulated by actors. *Perceptual and Motor Skills, 89,* 1195–1208.

Whiteside, S. P. (1999b). Note on voice and perturbation measures in simulated vocal emotions. *Perceptual and Motor Skills, 88,* 1219–1222.

Wiener, M., Devoe, S., Rubinow, S., & Geller, J. (1972). Nonverbal behavior and nonverbal communication. *Psychological Review, 79,* 185–214.

Williams, C. E., & Stevens, K. N. (1969). On determining the emotional state of pilots during flight: An exploratory study. *Aerospace Medicine, 40,* 1369–1372.

Williams, C. E., & Stevens, K. N. (1972). Emotions and speech: Some acoustical correlates. *Journal of the Acoustical Society of America, 52,* 1238–1250.

Wilson, E. O. (1975). *Sociobiology.* Cambridge, MA: Harvard University Press.

Wilson, G. D. (1994). *Psychology for performing artists. Butterflies and bouquets.* London: Jessica Kingsley.

Witvliet, C. V., & Vrana, S. R. (1996). The emotional impact of instrumental music on affect ratings, facial EMG, autonomic responses, and the startle reflex: Effects of valence and arousal. *Psychophysiology, 33*(Suppl. 1), 91.

Witvliet, C. V., Vrana, S. R., & Webb-Talmadge, N. (1998). In the mood: Emotion and facial expressions during and after instrumental music, and during an emotional inhibition task. *Psychophysiology, 35*(Suppl. 1), 88.

Wolfe, J. (2002). Speech and music, acoustics and coding, and what music might be for. In K. Stevens, D. Burnham, G. McPherson, E. Schubert, & J. Renwick (Eds.), *Proceedings of the 7th International Conference on Music Perception and Cognition, July 2002* [CD-ROM]. Adelaide, South Australia: Causal Productions.

Woody, R. H. (1997). Perceptibility of changes in piano tone articulation. *Psychomusicology, 16,* 102–109.

Zucker, L. (1946). Psychological aspects of speech-melody. *Journal of Social Psychology, 23,* 73–128.

*Zuckerman, M., Lipets, M. S., Koivumaki, J. H., & Rosenthal, R. (1975). Encoding and decoding nonverbal cues of emotion. *Journal of Personality and Social Psychology, 32,* 1068–1076.