

meanings (Wilson 1975). The ultrasonic squeaks of a young lemming denote alarm; the swinging steps and lifted tail of the male baboon summon his troop to follow; the "song" of the male white-crowned sparrow informs his fellows of his species, sex, local origin, personal identity, and readiness to breed or fight. Even the elaborate "dance" of the honeybee merely conveys information about the direction, distance, and quality of a nectar trove. But language can convey information about many more matters than these. In fact, it is the peculiar property of language to set no limit on the meanings it can carry. (pp. 208–209)

Let us now review the evidence for the informational content of animal signals. This review process will allow us not only to describe, for example, the honeybee's dance language, but also to make a stab at understanding why the specificity or sophistication of this system stops where it does.<sup>5</sup>

### 7.2.1 Information about Affective State

The pitch of the voice bears some relation to certain states of feeling. . . . a person complaining of ill treatment . . . almost always speaks in a high-pitched voice. (Darwin 1872, 88)

All organisms experience changes in affective state as a result of direct confrontations with the environment and, in some species, expected or imagined confrontations with the environment. The affective states I am referring to include both motivational and emotional. In a majority of species, affective states are responsible for the production of communicative signals. In this section I review what is known about the relationship between changes in affective state and the structure of communicative signals. This connection is particularly rich from a comparative perspective because it allows for an unambiguous comparison of species as diverse as red-winged blackbirds, chimpanzees, and human infants. As a reminder for those who have forgotten (!), humans begin life "talking" about their affective states and do so through laughter, crying, and growling. Only later do such emotive outpourings develop into full-blown speech.

On the basis of comparative observations, Darwin (1872) argued that animal vocalizations are designed to convey explicit information about emotional state. His logic went as follows:

1. Dominant individuals are generally larger (i.e., weigh more) than subordinates.
2. The pitch of an individual's voice is negatively correlated with its weight. Thus large individuals will have relatively lower pitched voices than small individuals.

5. In chapter 2, I pointed out that the early ethologists considered repertoire size to be limited by memory constraints.



3. Dominants will have lower pitched voices than subordinates.
4. Because of the relationship between size, pitch, and dominance, aggressive vocalizations will be low in pitch and submissive vocalizations will be relatively high in pitch.

Morton (1977, 1982) was the first to provide a comprehensive treatment of Darwin's prediction. Using data on avian and mammalian vocalizations, Morton provided qualitative support for the relationship between pitch and motivational state. Moreover, he extended Darwin's logic to other motivational states and acoustic features, laying out a set of "motivation-structure (MS) rules." For example, he predicted that aggressive vocalizations would be broadband and noisy, whereas fearful/appeasement vocalizations would be narrowband and tonal. Further, in contexts where mixed or ambiguous motivational states arise (e.g., alarm/mobbing calls), the acoustic structure of the signal should reflect this ambiguity (e.g., a chevron-shaped signal that rises and falls in frequency). Some of the patterns derived from MS rules are schematically illustrated in Figure 7.2.

More quantitative but comparatively narrower tests of Morton's MS rules have recently been carried out for several primate species. Analyses of a male

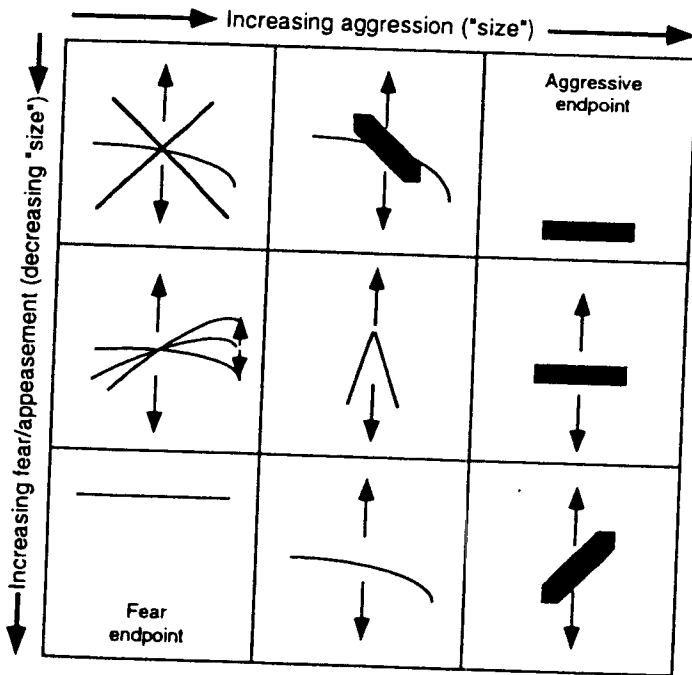


Figure 7.2

Some of the potential relationships between changes in motivational state and acoustic structure, emphasizing frequency (y-axis) and tonality/bandwidth. Thin lines represent tonal sounds; thick lines represent broadband noisy (atonal) sounds. Arrows indicate that frequency can shift up or down. The central panel represents a chevron-shaped sound, commonly associated with the context of alarm/mobbing—a mixed motivational state (redrawn from Morton 1982).

chimpanzee's vocalizations,<sup>6</sup> briefly discussed in chapter 4, indicated that the pitch of aggressive barks was relatively lower than the pitch of submissive screams (Bauer 1987). In contrast to this case study, results from research on two macaque species (Gouzoules and Gouzoules 1989; Gouzoules, Gouzoules, and Marler 1984) and the ring-tailed lemur (Macedonia 1990) failed to provide support for the predictions from MS rules. For example, both rhesus and pigtail macaques produce five acoustically distinct screams in the context of submission, each scream type conveying information about the type of aggressive interaction encountered (e.g., attacks involving dominant kin versus nonkin, and with or without physical contact). However, whereas one species uses a relatively low-frequency, broadband, atonal scream in the context of physical attack from a

6. Analyses were derived from a single film sequence.

dominant, the other species uses a relatively high-frequency, narrowband tonal scream. That is, similar motivational states appear to be associated with different acoustic signals in closely related species.

In addition to tests of the relationship between motivational state and acoustic structure, other studies have explored the relationship between body size and pitch. As previously stated (see also chapters 4 and 6), large males produce lower-pitched advertisement calls than small males in a number of anurans (Ryan 1988; Ryan and Brenowitz 1985). This correlation, however, does not hold in all species, including other anurans (Arak 1988; Doherty and Gerhardt 1984; Gerhardt et al. 1987) and nonanuran vertebrates (e.g., red deer: McComb 1991; humans: Künzel 1989).

Though it is important to understand why exceptions arise, Morton's MS rules, like Darwin's earlier predictions, were intended to explain quite general and comparatively broad patterns of communication. Recently, Hauser (1993a) used a data set of several hundred vocalizations from 43 nonhuman primate species (23 genera, including the prosimians, New and Old World monkeys, and apes) to examine, first, the relationship between body weight and pitch (a between-species comparison) and, second, the relationship between motivational state and pitch (a within-species comparison). In general, body weight accounted for a significant proportion of the variation in vocal pitch among species. Thus the largest species within the order Primates (e.g., gorillas, chimpanzees) produce the lowest-pitched calls, whereas the smallest species (e.g., bush babies, marmosets) tend to produce the highest-pitched calls. Important exceptions to this general pattern were, however, observed, indicating that for some genera, body weight is clearly not the most important factor underlying the range of frequencies used. Within the genus *Macaca*, for example, differences in the species-typical habitat may account for a significant proportion of the variation in frequency. Forest-dwelling species (e.g., lion-tailed macaques) use a narrow range of low-frequency calls (150–1000 Hz), whereas species living in more open habitats (e.g., bonnet macaques) use a broader range of low- and high-frequency calls (150–15,000 Hz). This explanation must be treated cautiously, however, since for some species, recordings were obtained from a captive setting, whereas for other species, recordings were obtained in a natural setting (see chapter 3 for details on how the environment can exert specific pressures on the design features of the signal). Nonetheless, the macaques represent a beautiful test case because differences in the range of frequencies used are unlikely to be due to species-specific differences in the mechanisms or hardware underlying phonation; though see Schön Ybarra (in press) for a review of subtle variation in vocal-fold structure among primates.

Results from analyses of the relationship between motivational state and pitch among nonhuman primates also corroborated the prediction set forth by Darwin and Morton. Across species, low-frequency calls were generally given during aggressive interactions, and high-frequency calls were given during submission or fear. As with the body weight–frequency relationship, however, not all species provided support for the predicted relationship between frequency and motivational state, and the overall relationship between motivational state and tonality was not statistically significant. For example, although most species within the Old World monkeys produced low-frequency vocalizations in the context of aggression, this was not consistently the case for the prosimians, New World monkeys, and apes. In summary, therefore, the pitch of the voice appears to convey approximate information about body weight and motivational state in many non-human primate species. These two factors, however, do not account for all, or even most, of the variation in vocal pitch. To tackle the issue of motivational-structural rules head-on, we will require more accurate measurements of individual body weights within species together with quantitative assays of motivational state. Concerning the latter, it will be useful if future researchers attempt to combine behavioral measures of motivational state (e.g., facial expressions and body postures associated with a particular context) with physiological measures (e.g., changes in cortisol or heart rate); see discussion of human facial expressions in section 7.3.3 for an elegant example of how behavioral and physiological measures can be combined.

Although a number of linguists have provided elegant analyses of how speech-like structures emerge in the developing child (Locke 1993; Oller 1986; Oller and Eilers 1992; Vihman 1986), relatively fewer studies have been conducted on the purely emotive sounds produced by young infants, with the exception of research on infant cries (Lester and Boukydis 1992; Zeskind and Collins 1987; Zeskind et al. 1985). Recently, however, a handful of papers have emerged focusing on the structure and function of gruntlike sounds (McCune and Vihman, in press; McCune et al., in press) and laughter (Nwokah et al. 1993). In all of these studies, the goal has been to provide a taxonomy of signal variation and stereotypy under different conditions and to assess the perceptual salience of various acoustic parameters. Analyses of the child's emotive utterances are of considerable significance with regard to comparative analyses and are discussed in the following paragraphs.

Biomedical researchers have studied crying as an assay for screening healthy and unhealthy infants (Colton and Steinschneider 1980; Lester et al. 1991). Developmentalists, in contrast, have analyzed the structure and contexts associated

with crying in order to understand how the child makes use of a limited communicative tool kit. As described in chapter 5, though infants, older children, and even adults cry, for the young preverbal infant crying represents one of the primary acoustic vehicles for transmitting information about affective state. Infant cries are also individually distinctive (J. A. Green and Gustafson 1983; Gustafson, Green, and Tomic 1984) and appear to be designed for long-range propagation (Gustafson, Green, and Cleland 1994).

From a functional perspective (chapter 6), it would make good sense for infant cries to convey at least some accurate information about changes in affective state and for caretakers to be sensitive to such covariation (Lester and Boukydis 1992; Lester and Zeskind 1978). Research has shown that infants produce at least two cry types, labeled *pain* and *hunger* cries. In general, parents are able to discriminate between these cries, and they show behavioral responses that are appropriate to the affective state conveyed. Infants may use cries to express other affective states, but thus far it has not been possible to assess whether the acoustic characteristics of these putative cry types are species-specific or represent individual or culture-specific changes due to subtle contextual influences.

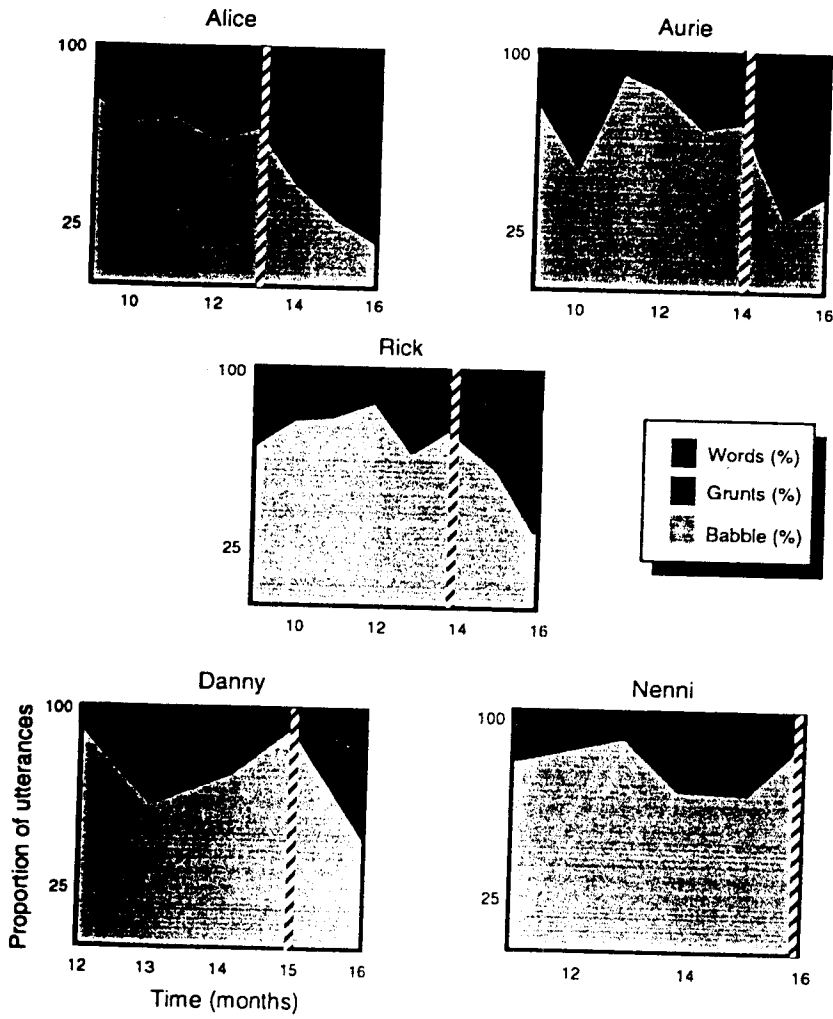
In a wide variety of mammalian species, individuals produce communicative grunts. Several authors, especially those working on nonhuman primates (Cheney and Seyfarth 1982; Stewart and Harcourt 1994), have argued that such vocalizations are not merely reflexive sound emissions but rather intentional signals that convey information about the caller's social environment (e.g., dominance relationships, group movement, intergroup encounters). Now, in a stimulating set of empirical and theoretical papers, McCune and colleagues (McCune and Vihman, in press; McCune et al., in press) have followed the approach of primatologists in an analysis of human infant grunts. Grunts were identified in infants on the basis of three shared criteria: "(1) abrupt glottal onset, (2) lack of supraglottal constriction, and (3) short duration" (p. 5). Given these acousticomotoric parameters, three grunt types were observed based on contextual information. "Effort grunts" appeared to be generated simply on the basis of the infant's effort to obtain or do something, such as stand up or reach an object. "Attention grunts" occurred in the absence of movement or effort, and appeared to be in response to an attentional interaction; no communicative intent could be observed. "Communicative grunts" were distinguished from the other two on the basis of gestures, including eye gaze, apparently designed to attract someone's attention, usually the mother; such attention-seeking utterances were typically produced when the infant was in need of help or wanted to show something off. Based on these operational definitions, observations of five infants were collected between

the ages of 9 and 16 months (Figure 7.3). In general (see summary, Figure 7.4), effort and attention grunts occupied a large proportion of the infant's early vocal repertoire. It was not until at least 13 months that communicative grunts appeared, and interestingly, there was a reasonably close correspondence between the onset of communicative grunts and the onset of referential word use.

The data presented in Figure 7.3 suggest that prior to the onset of communicative utterances, human infants produce a number of sounds that are mere expressions of changing affective states, including pain, hunger, and joy. Such sounds are also accompanied by facial expressions (see Lewis 1994). Eventually, however, some of these sounds are coopted for use in intentional communication, as occurs when infants use grunts to indicate their desires. Further support for the intentionality of such utterances comes from the observation that communicative grunts are commonly accompanied by nonvocal gestures, such as pointing to a desired object. Moreover, communicative grunts appear quite synchronously with the steep rise in the production of referential words, as well as what appear to be planned actions. A synthesis of this information, together with results on other communicative milestones, suggests an interesting ontogenetic system. I present this scenario (McCune and Vihman, *in press*) in Figure 7.4 and return to its implications later on in the chapter.

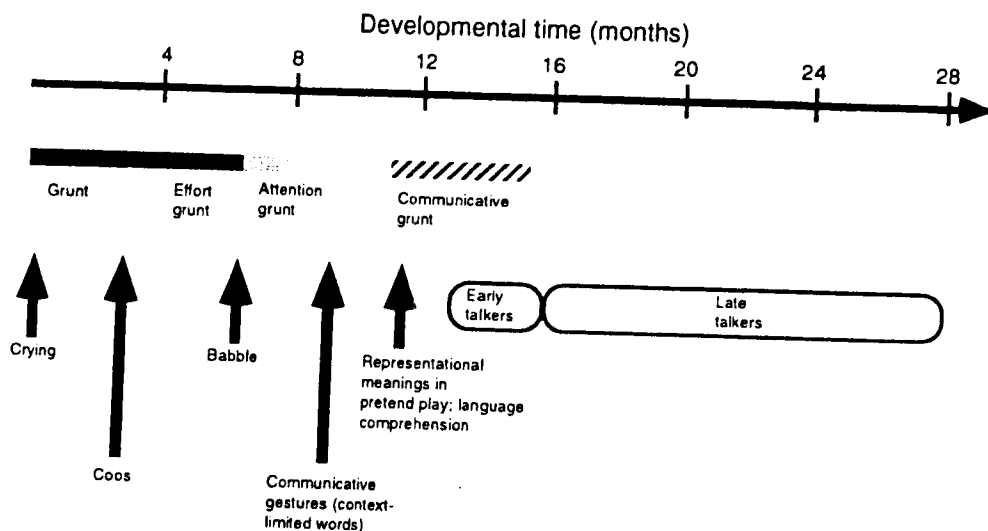
Research by Nwokah and colleagues (1993) has focused on the laughter of young children (three-year-olds) and, in particular, on documenting how changes in acoustic structure map on to changes in emotional state. Results suggest that children produce at least four different types of laughter: comment, chuckle, rhythmical, and squeal. The primary acoustic differences among these types were fundamental frequency contour, number of units, and unit duration. Such acoustic differences, in turn, appear to covary with the degree of excitement or arousal instantiated by the social context (in the current study, mother-child play). It will now be important to extend these sorts of analyses to younger children, different social contexts, different cultures, and perhaps, different species such as the great apes where laughter has been described (van Hoof 1972; for spectrograms, see Marler and Tenaza 1977). In addition, as laid out by McCune and colleagues for grunts, it will be interesting to establish when, developmentally, children begin to use laughs intentionally, as occurs when we snicker.

The literature reviewed thus far indicates that for nonhuman and human animals, acoustic signals encode some information about the signaler's affective state. Does the structure of the signal, however, have an effect on the affective state of relevant perceivers? McConnell (McConnell and Baylis 1985; McConnell 1990) has taken an interesting approach to this problem using cross-cultural data



**Figure 7.3** Variation in the onset of grunts, words, and babbles from six children ranging in age between 9 and 16 months. The striped vertical bar represents the age at which each child started producing communicative grunts (figure redrawn from McCune et al., in press).





**Figure 7.4**












A developmental scenario proposed by McCune and colleagues (*in press*) for the onset and change (structural and functional) of grunts by human infants, relative to other communicative signals and milestones. Early talkers are considered to be children who have referential words, based primarily on recorded productions of two consonants. Late talkers are those who use words and sentences.

on the sounds produced by humans to mediate the movement of trained pets (primarily dogs and horses), in addition to developmental experiments with young dogs. The logic of her argument is that if signals are designed to convey information about the sender's affective state, then such signals should have an effect on the affective state and subsequent response of perceivers. An important theoretical motivation for this view, discussed in chapters 2 and 6, was that selection should favor manipulative signalers (Dawkins and Krebs 1978; Krebs and Dawkins 1984).

Based on interviews with animal trainers and acoustic analyses of their signals,<sup>7</sup> McConnell found a reasonably close correlation between the acoustic structure of sounds used to arrest movement and those used to increase activity. In general, when trainers wanted their subjects to move, they produced multinote signals that were short and frequency modulated. When they wanted their subjects to stop moving, they produced single notes that were relatively long in duration.

7. The sample was obtained from trainers speaking the following native languages: English, Spanish, Swedish, High German, Polish, Basque, Peruvian Quechua, Finnish, Chinese, Korean, Arabic, Farsi, Pakistani, Navajo, Shoshone, and Arapaho.

**Table 7.1**  
Relationship between Whistle Acoustic Structure Produced by Shepherds and Intended Message for Shepherd's Dog (redrawn from McConnell and Baylis 1985)

	Multiple Notes				Single Notes						
											
Fetch	8	3	3	0	0	0	0	0	0	0	0
Stop	0	1	0	0	4	2	6	0	1	0	0
Go left	2	3	0	1	3	0	1	1	0	2	1
Go right	2	0	1	1	1	0	1	0	4	4	0

Cell values represent the number of signals that were assigned to the context on the left by the author (McConnell and Baylis 1985). The thick vertical line divides multiple- from single-note whistles.

Some examples are provided in Table 7.1 based on analyses of shepherds' whistles to dogs (McConnell and Baylis 1985).

In experiments with young dogs, McConnell (1990) trained one group of individuals to approach in response to a series of short, rising-frequency notes and stop-sit in response to a long, descending-frequency note; a second group was trained on the same signals, but with motor responses reversed. Results showed that the short, rising-frequency notes more readily caused individuals to approach than a single descending-frequency note. Although the long descending frequency caused more individuals to sit-stay, this result was not statistically significant.

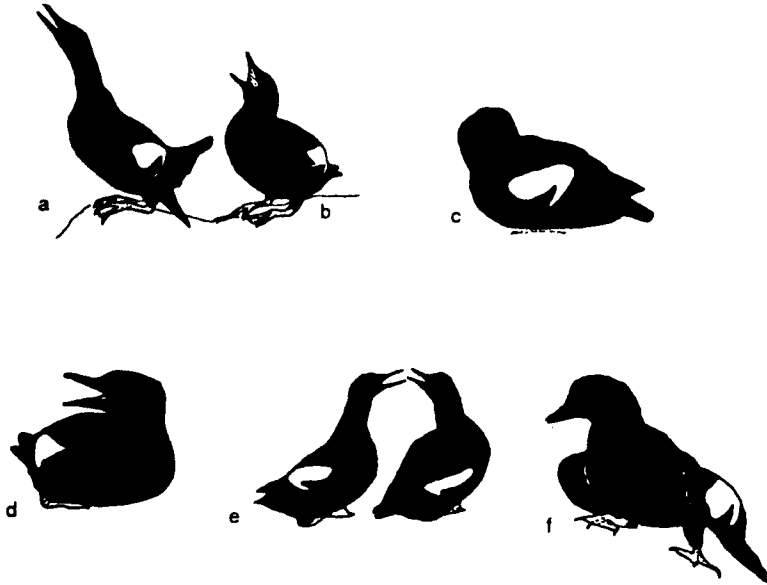
The issues raised by McConnell's work are also directly relevant to studies of intraspecific communication, and in fact such studies have already sneaked into some of our earlier discussions. Specifically, in chapter 2, I reviewed the controversy between what some have called the traditional ethological view and the behavioral ecological view (Caryl 1979, 1982; Hinde 1981). To reiterate, much of the early ethological literature suggested that animals convey specific information about their emotional/motivational state (i.e., about the probability of initiating a particular behavioral response), and such information is used to select fitness-maximizing responses. Behavioral ecologists criticized this perspective, arguing instead that honest signaling of motivational state was evolutionarily unstable, open to cheaters who escalate dishonestly. Consequently, those working within

the framework of behavioral ecology tended to argue that signals are designed to manipulate the behavior of the perceiver, and as such, are not informative with respect to motivational state.<sup>8</sup> These two views have made some peace with each other, based on both more recent theoretical models and empirical findings. In particular, although circumstances exist where individuals will be favored to conceal information about their motivations (e.g., Maynard Smith's "war of attrition" model makes this prediction for situations where the only cost incurred by individuals is time spent in an aggressive encounter), other studies suggest that accurate information about underlying motivations will be favored. Van Rhijn and Vodegel's (1980) model suggests, for example, that in species with individual recognition and repeated interactions among group members, truthful signaling about motivational state (subsequent behavior) will be selected for. Given these distinctions, let us look at a few studies where detailed analyses of the relationship between signal structure and motivational state have been explored.

One of the strengths of early ethological research was its attention to behavioral detail, especially fine-grained analyses of the motor actions involved in a display and the sequences of such displays used in communicative interactions. The general focus was on aggressive competition, especially in birds and fish. Most of the studies provided data on three components of the interaction: (1) type of display used by the actor, (2) response to the display by a reactor, and (3) follow-up behavior by the original actor. Thus, in a classic analysis format, there would be information on the probability of aggressive attack by the actor given that the reactor either attacks or flees. What is clear from these early studies is that *some* information about motivational state is conveyed, but as Caryl's (1979) reanalysis suggests, such information is not *absolutely* reliable. In other words, if A receives an aggressive threat display from B, A can not be sure that B will attack even if A holds his ground instead of running away.

An elegant and extremely thorough treatment of this issue was presented by Nelson (1984) based on observations of the pigeon guillemot (*Cephus columba*), a seabird. This species performs a variety of visual and vocal displays during territorial disputes (Figure 7.5). In general, territorial males consistently win aggres-

8. Recall, however, that the behavioral ecologists did not deny the existence of information exchange during communication. They suggested that individuals could obtain accurate information about such things as fighting ability by directly observing a competitor's resource holding potential (RHP). Attributes associated with RHP included body size and various ornamental traits such as antlers, tails, canines, etc. Moreover, in Krebs and Dawkins' (1984) revised framework, cooperative communicative interactions were seen as precisely those where information is exchanged by conspiratorial whispers.



**Figure 7.5**  
The aggressive displays of the pigeon guillemot used during territorial disputes. (a) neck-stretch, (b) hunch-whistle, (c) bill-tuck, (d) trill, (e) duet trill, (f) trill-waggle (from Nelson 1984).

sive encounters against intruders, and most resolutions arise in the absence of fighting. In fact, based on more than 600 aggressive encounters, Nelson never observed an injury. The owner's use of the neck-stretch display was a highly effective threat as evidenced by the fact that intruders flew away in response to this display on 43% of all occurrences. In contrast, displays used by the owner failed to provide an accurate indication of the probability of attack: only 14% or less of all display types were followed by attack. Last, Nelson was able to use sophisticated statistical techniques to accurately predict the probability of the owner's subsequent behavior based on the spectrotemporal properties of the hunch-whistle signal. Specifically, owners were more likely to sit than attack when note duration or frequency increased; owners were more likely to move than attack when note duration decreased. These data indicate that contrary to Caryl's assertions, some of the pigeon guillemot's displays convey quite accurate information about subsequent behavior, information that is clearly used by individuals in selecting an appropriate response.

Countercalling, either in the form of one-on-one interactions or vocal choruses (reviewed in M. D. Greenfield 1994), is common among nonhuman animals and forms the basis of the human child's entry into the world of acknowledged

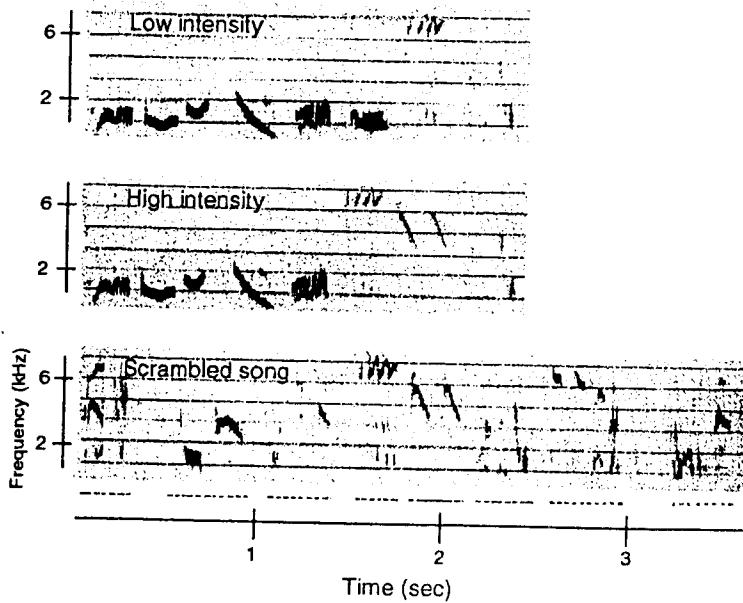
communicative competence (i.e., being able to interact conversationally).<sup>9</sup> During such vocal duels, individuals may obtain information about their opponent's motivational state, and in nonhuman animals, such information is especially likely to be proffered when actor and reactor are territorial neighbors (i.e., familiar and engaged in frequent interactions).

In the European blackbird (*Turdus merula*), territorial males sing during defensive interactions with intruders and while attempting to attract mates. As noted for other species, much of the acoustic variation in song structure covaries with variation among individuals. However, as Dabelsteen and Pedersen (1990; Dabelsteen 1992) have documented, a considerable proportion of the variation is accounted for by changes in the motivational state<sup>10</sup> of the singer, changes that occur rapidly during vocal exchanges (Figure 7.6). Thus, at the start of a vocal exchange, males typically begin with what are called "low-intensity" songs. Once a response is elicited, the male rapidly moves into a "high-intensity" song. Thus, high-intensity songs appear to be indicative of an interaction that has escalated. The final level of intensity, and apparently, aggressive competition or threat, occurs when a rival male enters the resident's territory. At this point, males begin singing a "scrambled song." Dabelsteen's observations indicate that in the heat of a vocal duel, males rapidly alternate between these song types, and interject two other types as well. Sometimes, however, males match song type for song type; such song matching is common among oscine song birds and appears to provide some information about competitive ability (Kroodsma 1979; Todt 1981). Though the precise function of different patterns of vocal exchange is not yet clear, the European blackbird system provides a good example of a graded system of communication (see chapter 3), with song variants mapping onto changes in motivational state.

The studies on human infants reviewed earlier indicated that before they begin to communicate about objects and events in the external environment, infants convey information about their affective state. Human adults also communicate nonlinguistically about affective state, and one of the primary arenas for such communication occurs in infant-directed speech (see chapter 5). These signals, known in linguistic circles as the prosodic, paralinguistic, or melodic features of

9. In fact, one of the most striking hints that a child may be suffering from a developmental problem, as commonly occurs among young autistic children (Frith 1989), is that he or she fails to engage in conversational interactions, including the maintenance of vocal and visual attention.

10. Note that in Dabelsteen's work, as in much other research in this area, no independent measure of motivational state has been provided. Rather, motivational state is assumed on the basis of behavioral state.



**Figure 7.6**

Three songs produced at different intensities by European blackbirds. Low-intensity songs are typically heard at the start of a song bout and are then followed by high-intensity songs, especially when the initial singer hears a vocal response from a neighbor. When a rival enters the singer's territory, he often switches to a scrambled song (highest intensity) (redrawn from Dabelsteen 1992).

language, not only provide the infant with direct information about its caregiver's emotional state (Fernald 1992a, 1992b), but also, in some circumstances, provide crucial information about the syntactical structure of language (Bahrack and Pickens 1988; Fernald and McRoberts 1994; Jusczyk et al. 1993; for some exceptions, see Gerken, Jusczyk, and Mandel 1994).

Adult speech to other mature listeners of the linguistic community also carries important information about the speaker's affective state. Studies of the acoustic correlates of affective state have, however, played a relatively minor role in the study of human communication (Goldbeck, Tolkmitt, and Scherer 1988; Hess, Scherer, and Kappas 1988; Scherer 1986; Scherer and Ekman 1982). Moreover, most of the research has been carried out with actors or public speakers rather than with naturally occurring conversations or utterances. Nonetheless, some general patterns have emerged, especially in terms of the relationship between relative pitch of the voice (fundamental frequency) and emotional state (Table 7.2). For example, Tartter (1980) has shown that smiling causes an increase in pitch, whereas frowning causes a decrease (Tartter and Braun 1994). Listeners

**Table 7.2**  
Some of the Acoustic Correlates of Affective State (from Scherer 1986)

Emotion	Pitch Level	Pitch Range	Pitch Variability	Loudness	Tempo
Happiness/joy	High	?	Large	Loud	Fast
Confidence	High	?	?	Loud	Fast
Anger	High	Wide	Large	?	Fast
Fear	High	Wide	Large	?	Fast
Indifference	Low	Narrow	Small	?	Fast
Contempt	Low	Wide	?	Loud	Slow
Boredom	Low	Narrow	?	Soft	Slow
Grief/sadness	Low	Narrow	Small	Soft	Slow
Evaluation	?	?	?	Loud	?
Activation	High	Wide	?	Loud	Fast
Potency	?	?	?	Loud	?

find both acoustic changes to be perceptually meaningful (i.e., indicative of the speaker's emotional state).

The reported correlations between voice characteristics and emotional state suggest that a person's speech provides a window into what he or she feels. Given that one can voluntarily manipulate such acoustic features, however, it seems likely that listeners would use more than the voice to evaluate signal honesty. Indeed, Ekman's work on deception suggests this kind of multimodal approach to signal perception (see section 7.4).

Thus far, our discussion of the relationship between affective state and signal structure has focused on the auditory domain. Similar issues also arise within the visual domain and form the core of current debates over the function of facial expressions. One view, espoused by Ekman and his colleagues (Ekman 1973, 1992; Ekman and Friesen 1969, 1975; Ekman, Levenson, and Friesen 1983), suggests that human facial expressions reflect the signaler's emotional state. A number of such expressions are universal, based on significant physiological changes (Figure 7.7). The opposing view, defended by Fridlund (1994) and J. A. Russell (1994), is much more Dawkinsian-Krebsian (Dawkins and Krebs 1978; Krebs and Dawkins 1984), suggesting that facial expressions function to manipulate the emotional state and behavior of others and that the evidence for universality is weak. As in the animal literature, both views clearly share a piece of the explanatory pie.

For more than twenty years, Ekman has engaged in a rigorous research program designed both to quantify the neurophysiological and anatomical substrates



**Figure 7.7**

A sample of facial expressions from Duchenne's (in press) classic work. In frame 31, stimulation of the zygomatic major muscle leads to what Ekman (1989) has called the Duchenne smile, or the smile of enjoyment.

underlying facial expressions and to assess whether universal expressions exist, and if so, why. A particularly attractive feature of this work is that it lays out a set of methodologies that are quantitative and, further, that appear to be primed for importing into studies of nonhuman animal expressions.

As I have pointed out, the most heated debate in the literature on facial expressions concerns the evidence for cross-cultural universals. Russell claims that there is very little evidence for universals. Ekman claims that the evidence is overwhelmingly in favor of universals and that Russell has simply misinterpreted or misrepresented the existing data. The core of this debate emerges from differences in the definition of *universals* and the methodological approach that is best suited for testing universality. Specifically, Russell argues that a universal facial expression must show no variation in either production or perception across cultures, whereas Ekman argues that some variation is to be expected but that



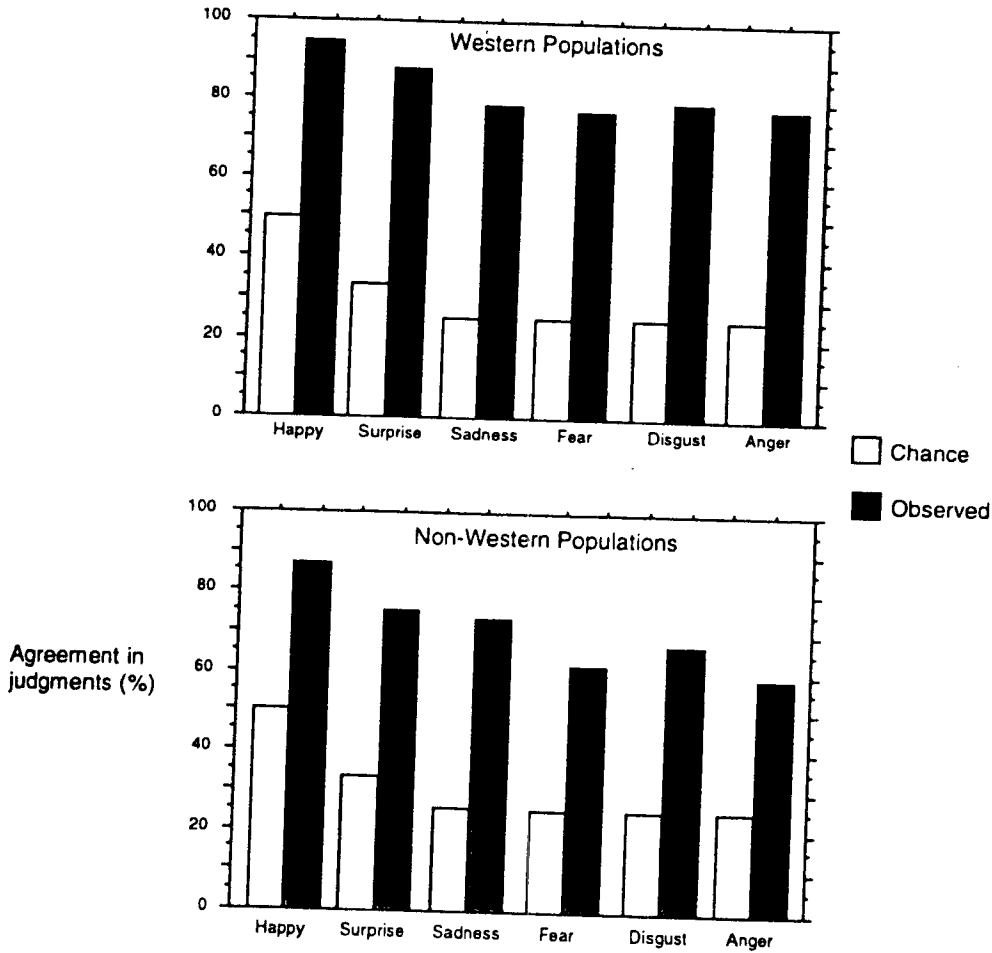
this in no way injures the claim for universality—a trait shared in the same basic form by most humans. Whereas Russell's detailed review suggests that there is little consensus on facial expression-emotion correlations, Ekman's reanalysis of the data presented by Russell provides a quite different picture, arguing strongly for universality. I present this reanalysis in graphical form in Figure 7.8.

In addition to the general findings already reported, a recent paper by Levenson and colleagues (1992) provides some of the strongest support for the universalist's position, based on behavioral and physiological analyses of expressions in Americans and Minangkabau. The Minangkabau are a group of people living in Western Sumatra. The society is Muslim, matrilineal, and agrarian. Most important for the current discussion, however, is the fact that overt displays of negative emotion are strongly discouraged.

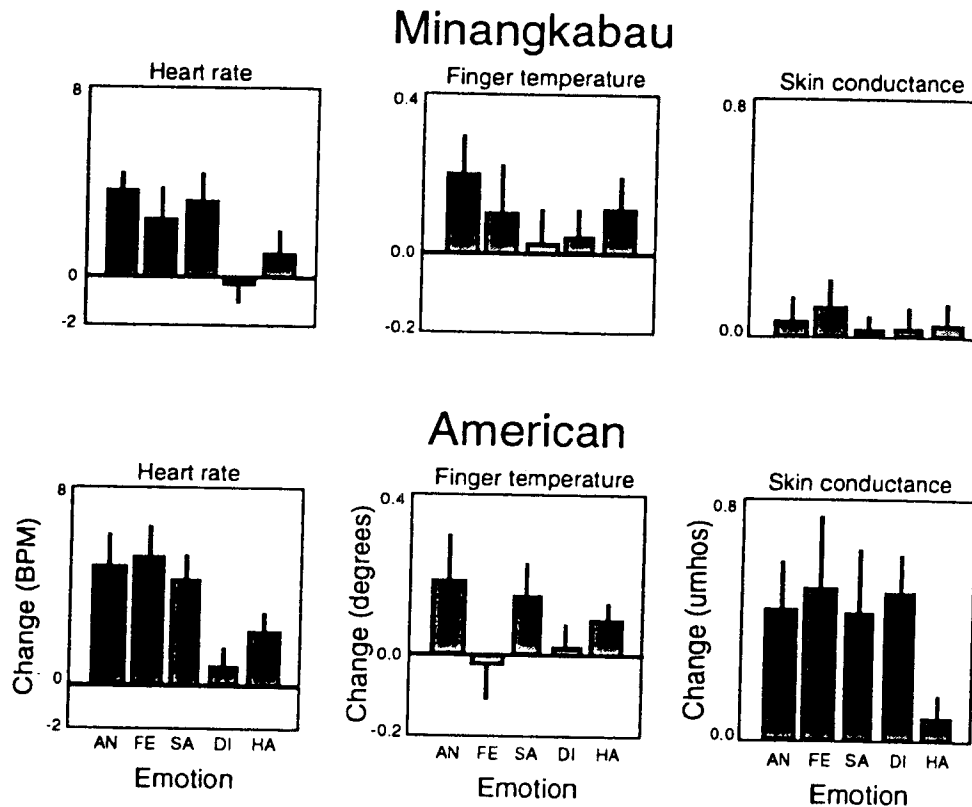
To test for cross-cultural coherence within emotions, as well as consistent differences across emotions, Levenson et al. carried out a directed facial action task. In this task, subjects are required to move particular facial muscles based on instructions read by an experimenter; some subjects carried out these instructions while looking in the mirror (to obtain feedback), whereas others did so in the absence of mirrors. While carrying out the instructions, subjects were monitored for a host of physiological changes, including heart rate, finger temperature, skin conductance, finger pulse transmission and amplitude, and respiratory period and depth; these factors were measured in order to provide a comprehensive description of autonomic nervous system (ANS) reactivity.<sup>11</sup> Results, some of which are presented in Figure 7.9, generally indicated significant differences in ANS responses to different emotions and no significant difference within emotions across cultures. These data strongly suggest that ANS activation from facial expression is consistent cross-culturally.

In the description of the Minangkabau, I alluded to a societal convention regarding overt expression. Specifically, overt, public expressions of negative affect are strictly frowned upon. This social impact on communicative expression is reminiscent of Marler's (Marler, Dufty, and Pickert 1986a; Marler, Karakashian, and Gyger 1991) notion of an audience effect (see summary of work in the subsection "Referential Signals in the Domestic Chicken") and has been explicitly examined by Fridlund and colleagues (Chovil 1991; Fridlund et al. 1990; Fridlund, Kenworthy, and Jaffe 1992) with regard to the social dynamics of human facial expression. In particular, Fridlund's work challenges the facial-expression-as-

11. Previous research had demonstrated that such factors reliably predicted emotional differences among American subjects (Levenson, Ekman, and Friesen 1990; Levenson et al. 1991).



**Figure 7.8** Plot of data tabulated by Ekman (1994) on cross-cultural agreement in facial expression, comparing Western (top) and non-Western (bottom) populations. Values expected by chance (white) were derived on the basis of the following logic. Happy is set at 50%, based on the fact that the choice is between either a positive or negative emotion. Surprise is set at 33.33%, based on the empirical finding that the expression is most often identified as either surprise, fear, or happiness. The four negative emotions, sadness, fear, disgust, and anger, are set at 25%, based on the probability of picking one of these four terms (for a more detailed description, see Ekman 1994, 272).



**Figure 7.9**  
Data from Levenson and colleagues (1992) showing the similarity in autonomic response to different facial expressions. AN: anger; FE: fear; SA: sadness; DI: disgust; HA: happiness.

emotional-readout view (*sensu* Ekman) by making two points: first, all facial expressions are enacted *to* some social group, whether present or imagined; second, from a behavioral ecology perspective, such expressions have been designed to manipulate others rather than to provide them with explicit details of their affective states (i.e., expressions are highly deceptive). This view does not deny the relationship between emotion and facial expression. Rather, it argues that the relationship is more complex than previously entertained. What is the evidence for this alternative perspective?

To test the idea that all facial expressions are mediated by a social context, even when a person is physically alone (e.g., in a room that is visually and auditorily isolated), Fridlund and his colleagues (Fridlund et al. 1990; Fridlund, Kenworthy, and Jaffe 1992) have taken advantage of techniques in mental imag-

ery designed by Kosslyn (reviewed in Kosslyn 1994). In one task, a subject arrives at the testing room either alone or with a close friend and is then placed in one of four conditions while viewing a comical videotape: (1) arrives alone and views tape alone; (2) arrives with friend and views tape alone but is told that friend is being tested on other sorts of psychological tests; (3) views tape alone but is told that friend is watching the same tape in another room; and (4) views tape with friend in same room. Results indicate a significant increase in the extent of smiling (quantified by EMG records) from conditions (1) to (4), although, interestingly, there were no differences between conditions (3) and (4);<sup>2</sup> the lack of difference between these two conditions suggests that imagined and actual presence of a friend *can* exert comparable effects on at least one form of facial expression, smiling.

From these kinds of results, Fridlund (1994) has argued that there are at least two problems with viewing facial expressions as simply mirrors onto the emotions. First, he claims that there is no agreed-upon definition of emotion, and thus the phenomenon to be explained is far too elusive for consensus, at least at present. In support of this claim, he quotes from a variety of sources, highlighting the conflicting perspectives. Second, and more critically, Fridlund argues that although current studies show a significant correlation between ANS change and facial expression (either through the directed facial action task or by means of recalling a particularly emotional event), they fail to tease apart two additional factors, believed to be critical to the face-emotion view. Specifically, physiological changes must be associated with something more than the subject's general state of arousal (e.g., increased heart rate associated with anger relative to sadness because the former requires greater energy) and must be the result of changes in emotion rather than in display function.

In summary, whereas Ekman explicitly argues in favor of the emotional content of facial expressions, maintaining that they have been designed to convey such information, Fridlund explicitly denounces this view, arguing instead that facial expressions are manipulative communication displays and can "be understood without recourse to emotions or emotion terms" (p. 186). My own reading of this literature is that the debate actually confuses levels of analysis. Whereas Ekman's work has generally focused on the *mechanisms* underlying facial expression (e.g., changes in physiology, brain state), Fridlund has considered the *function* of facial

2. Note that in Fridlund's (1994) review of this work he states that there was "a monotonic increase in smiling from" (p. 162) conditions 3 to 4. From his Figure 8.8 (p. 162), however, these two conditions have means and standard deviations that are virtually identical.

expression. As I have repeatedly pointed out in this book, both perspectives are important, but they are not mutually exclusive.

### 7.2.2 Information about the External Environment

**Honeybees: A Puzzling Case?** One of the nice properties of human language is that if I am hungry and want the address of a good restaurant, I can stop someone on the street and ask for directions. If the person is familiar with the area, he or she might be able to say something like

There's this wonderful little French restaurant about 10 minutes away. Go to the first light and turn right, then walk about 10 blocks until you see a shoe store on the left side of the street. Turn left and walk down the short alley. The restaurant, called *La Vie et La Bouf*, has a red and white front, and a large picture glass window filled with bottles of Bordeaux, columns of salami, and baskets of baguettes.

These instructions provide explicit information about the location of the restaurant, how far away it is, and its quality—at least based on the experience of the person providing the information. Humans looking in at the nonhuman animal world would surely turn and say, "Beat that! No animal can communicate with such expressive power." In my view, they would be partially wrong.

