

CHAPTER ONE

Prelinguistic Vocalizations

Among the most intriguing aspects of human language is its sound system and how the human infant comes to acquire it. Three major questions are addressed in this chapter. First, are there any similarities to be found between features of animal vocalizations and those of early infant vocalizations? In other words, can possible phylogenetic origins of human infant vocalizations be found in the vocalizations of nonhuman primates? If so, where exactly is the overlap? Alternatively, are even the earliest stages of human infant vocalizations quite distinct from animal vocalizations? Second, what are the major developments in the early ontogeny of the human vocalization system, and do they vary across infants with different characteristics or backgrounds? The focus here is on such features of vocal development as the onset of canonical (reduplicated) babbling, phonetic preferences, consonantal repertoire, consonantality (degree of consonant use), and complexity (combination of different consonants). Third, do variations in features of babbling across infants make a difference in language acquisition? Is just babbling itself an important precursor to language, or is the quality of babbling also important?

In this chapter, infant prelinguistic vocalizations are considered to be all phonated sounds (with vibration of the vocal cords) that are audible and are *not* crying, fussing, laughing recognizable words, imitated animal sounds, or imitated conventionalized expressions (*uh-oh*). These criteria are consistent with those used in most research on early infant vocalizations except that some researchers exclude also grunts (for example, Oller & Lynch, 1992) and some include words (for example, Vihman & Greenlee, 1987). Grunts, even those that are simple vegetative effort sounds, need to be included because they have been emphasized in phylogenetic continuity (McCune, Vihman, Roug-

Hellichius, Delery, & Gogate, 1996, see below). For some infants, they have early situational meaning (see chapter 2). Treating prelinguistic vocalizations as a separate category from adult-modeled words or expressions allows the relationship between them to be more easily assessed. The distinction between the two categories is not always an easy one, however (compare also Vihman, Macken, Miller, Simmons, & Miller, 1985).

This chapter treats only the phonetic aspects of vocalizations. Mapping of vocalizations onto context, or the origin of meaning, is the subject of the next chapter. The relationship between vocalizations and gestures is discussed in chapter 3.

Phylogenetic Origins of Human Infant Vocalizations

Unlike the case of communicative gestures, as is apparent in chapter 3, the evolutionary roots of human vocalizations are difficult to trace. Nonhuman primates emit a number of sounds that have been categorized, often with the aid of sound spectrograms, into barks, grunts, roars, screams or screeches, howls, squeals, growls, chatters, hoots, and pants. These categories are sometimes subdivided into different types, such as waa barks versus shrill barks. (Laughter, whines, and whimpers are not included here because they are eliminated from our definition of human infant vocalizations.) The categories appear to be quite similar across several species, for example, baboons (Hall & DeVore, 1965), langurs (Jay, 1965), chimpanzees (Goodall, 1965, 1986; Reynolds & Reynolds, 1965), mountain gorillas (Fossey, 1972; Schaller, 1965), orangutans (MacKinnon, 1974), and pygmy chimpanzees (bonobos) (Mori, 1983; De Waal, 1988). But not all categories are used by each species. Marler (1976) compared acoustic aspects of chimpanzee vocalizations from his recordings on the Gombe Reserve in Tanzania with spectrograms of gorilla vocalizations provided by Fossey (1972). By merging acoustically similar calls while ignoring their contexts and the characteristics of the sender, Marler reduced Van Lawick-Goodall's (1968) original 24 categories for the chimpanzee to 13, all of which overlapped with Fossey's vocalization categories for the mountain gorilla (Table 1–1). Fossey reported 3 additional categories, roar, growl, and a sound idiosyncratic to a single individual. Reynolds and Reynolds (1965) reported roar and growl for the chimpanzees of the Budongo Forest, but apparently the chimpanzees at Gombe do not use these calls. Mori (1983) suggested that *roar* should

Table 1–1. Morphological Equivalents in Vocalization of the Chimpanzee and Gorilla

| | Chimpanzee Call | Probable Gorilla Equivalent |
|-----|-----------------|--|
| A1 | Pant-hoot | Hoot series |
| A2 | Laughter | Chuckles |
| A3 | Scream | Scream |
| A4 | Rough grunt | Belch |
| A5 | Pant | Copulatory pant |
| B6 | Squeak | Cries (1) |
| B7 | Whimper | Cries (2) (also whine ?) |
| B8 | Waa-bark | Wraagh (1) (short form) |
| B9 | Wraaa | Wraagh (2) (long form) |
| B10 | Grunt | Pig grunt (1) (given in train) |
| B11 | Cough | Pig grunt (2) (given singly) |
| C12 | Pant-grunt? | Pant series? |
| C13 | Bark? | Hoot-bark? (Also hiccup-bark and question-bark?) |

Note: The correspondence is deemed most reliable in pairs labeled A and least in those labeled C.

Source: From P. Marler (1976), *Social organization, communication, and graded signals: The chimpanzee and the gorilla*. In P. P. G. Bateson & R. A. Hinde (Eds.), *Growing points in ethology* (p. 246). London: Cambridge University Press.

Reprinted with permission from Cambridge University Press.

be merged with *wraah*, and then the same vocalization is common to gorillas, chimpanzees, and bonobos.

An interesting difference between vocalizations in gorillas and chimpanzees, which was highlighted by Marler (1976), is that most call categories were found across age and sex groups in the Gombe chimpanzees but restricted to the silverback male in Fossey's gorillas. High-ranking male chimpanzees do engage in one particular type of vocalization, the pant-hoot, more than do low-ranking chimpanzees, and the alpha male responds less to the pant-hoot of other males (Mitani & Nishida, 1993). In general, in the intergroup communication of most primate species, "the loud calls are a male attribute" (Deputte, 1982, p. 68). Female chimpanzees also have a different pant-hoot from males in that it lacks a "climax" section; in addition, they bark more than males (Marler, 1976). It is claimed in later research (Harcourt, Stewart, & Hauser, 1993) that Fossey's (1972)

finding of vocalization dominance by the silverback gorilla was due to lack of habituation of the animals in her study. Although Fossey admitted this problem for the early period of her study, data were collected over 40 months, so it seems unlikely that all of her findings are distorted by reactivity. Nevertheless, Harcourt et al. (1993) did find more distribution of vocalizations across the sexes. At least close contact vocalizations, such as the grunts that are predominant in gorillas, were more widely used, although adult males still grunted the most. Such vocalizations were not heard in infants younger than 1.5 years, and they increased with age. The majority occurred as part of a vocal exchange within 1 second of each other. Lone gorillas (always male) were quite silent (See also Schaller, 1965; Stewart & Harcourt, 1994).

Harcourt et al. (1993), unlike Fossey (1972), did not record any roars or growls in their observations of gorillas. The frequency of “close” gorilla calls was about 8 per hour for adults, whereas Marler reported a rate of 10 to 100 per hour for chimpanzees. The chimpanzee rate is inflated because it includes all types of vocalizations and is based on recordings done at a feeding station where vocalizing is typically higher. Nevertheless, the comparison reflects a real difference in vocalizing tendencies between the species. Although chimpanzees remain silent for long periods, particularly during patrol and consortship, they can also be extremely vocal, even engaging in choruses of calling that resemble singing (Goodall, 1986). Bonobos also engage in “contest hooting,” but this appears to be a more rapid vocal dialogue (De Waal, 1988).

Although researchers have been able to formulate discrete categories for nonhuman primate vocalizations, they stress that by comparison with birdsong, vocalizations of monkeys and apes are variable and grade into each other. In the graded vocal repertoire of the chimpanzee, the fundamental frequency and duration of calls also vary independently (Marler, 1975). Many years ago, Marler (1965) pointed to these characteristics as presaging the human vocalization system, which is also variable and continuous even though humans impose discrete sound categories on their perception of the acoustic signal. Nonhuman primates also impose categorical boundaries in that they do not respond to a call whose duration exceeds its norm in call production (Snowdon, 1982). Marler (1975, 1976) has suggested that only non-territorial, multimale primate groups that communicate over short distances on the forest floor can use graded sounds. Such characteristics are consistent with speculations about the social organization of early

humans (Marler, 1975). Discrete, species-typic signals are also less essential in an environment where the sounds of other species rarely intrude to cause confusion. The danger of alien sound intrusion is minimal “when a species is living out of earshot of organisms similar in size and structure and when sounds are used at sufficiently close range that visual or other cues can confirm the identity of the signaling animal” (Marler, 1965, p. 565). Pure, discrete sounds, such as hoots, are still used by nonhuman primates to communicate over long distances. In contrast, birdsong is composed of stereotypic sounds that are discontinuous, very unlike human vocalizations, in part because each species competes with others for the sound space. More recently, however, describing birdsong as discrete and primate calls as graded has been termed an oversimplification (Marler & Mitani, 1988). Newman and Symmes (1982) argued that in fact, the graded nature of primate vocalizations is found only in the young and that these vocalizations become more discrete and less variable with maturity.

There appears to have been early general agreement that birds are uniquely similar to humans in exhibiting vocal learning (Nottebohm, 1975). Absence of vocal learning in apes was supported by the absence of local dialects in vocalizations (Marler & Tenaza, 1977). Snowdon (1982) did report subspecific differences in the long calls of saddleback tamarin monkeys living in different areas of the Amazon, but this finding is complicated by the possibility of interbreeding among subspecies. Snowdon (1982) also reported that pygmy marmosets that lost their parents never attained a fully adult vocal repertoire; this seems to be stronger evidence for vocal learning. In addition, Mitani and Gros-Louis (1998) recently documented vocal accommodation between adult male chimpanzees in chorusing – that is, calling bouts of overlapping pant-hoots. Two male pairs were found to produce calls when chorusing together that were acoustically more similar than calls they made with other males. One male’s chorused calls were closer acoustically to those his partner produced alone than to his own calls produced alone. Thus, this single chimpanzee provides preliminary evidence that chimpanzees are able to alter their calls during choruses of pant-hoots to match another’s calls.

The clearest link between human and nonhuman primate vocalizations appears to lie in the prosodic features displayed by nonhuman primate signals, namely pitch, timing, and intensity. These three features are physically described as follows: “the fundamental frequency contours, which give a language its characteristic melody; the duration

. . . measures, which give a language its characteristic rhythm; and the amplitude patterns, which give a language its characteristic patterns of loud versus soft syllables” (Levitt, 1993, p. 385). Pitch variation in nonhuman primate signals can be seen in the waa–barks of chimpanzees (wiew–barks of bonobos), the melodious “legato hooting” of bonobos (De Waal, 1988; Marler & Tenaza, 1977), the intergroup spacing calls of cotton-top tamarins with energy distribution across several formants (Snowdon, 1982), the long-distance J-calls of pygmy marmosets (Snowdon, 1982), and the harmonically structured clear calls of macaques (Brown, 1982). Such frequency modulation, along with repetition, makes calls more localizable (Snowdon, 1982). Prosody or the musical quality of speech is sometimes proposed as the earliest form of hominid vocal communication, and the roots of pitch contours associated with human emotions may be quite ancient, evolutionarily speaking (Hauser, 1996). In birds and mammals, high-pitched vocalizations tend to be associated with either fear or affiliation, and low-pitched, with aggression (Hauser, 1996). Darwin (1877) suggested that early hominids’ first form of vocal communication was expressed in song, often for emotive purposes in courtship (see review in Donald, 1991). Studdert-Kennedy (1991) speculated that prosody “perhaps first followed an independent course of evolution, to be modified and integrated into the linguistic system only as longer utterances and more finely differentiated syntactic functions emerged” (p.9). Prosody is also, as we shall see shortly, among the earliest developments in human ontogeny; and Ferguson and Macken (1983) have suggested that the “ontogenetic primacy of prosodic phenomena may reflect a phylogenetic primacy” (p. 238).

Differences between the Nonhuman Primate and Human Vocal Tracts

According to Bastian (1965), monkeys and apes differ from humans in the simplicity and steadiness of upper vocal tract configurations produced by their auditory signals. Human linguistic signals are distinguished by the “incessant occlusions of the vocal tract” (p. 595) to produce transient resonant patterns. Thus, although all mammals possess respiratory muscles that blow air through the vocal tract, features of the human upper vocal tract allow articulatory capabilities beyond the reach of nonhuman primates. “The open-to-closed and closed-to-open articulatory actions may occur at various places in the upper

Cambridge University Press

978-0-521-59299-4 - Routes to Child Language: Evolutionary and Developmental Precursors
Joanna Blake

Excerpt

[More information](#)*Prelinguistic Vocalizations*

7

vocal tract; they may be partial or complete; they may or may not be accompanied by concurrent glottal action; and, most important, they may be readily combined in many different sequences” (Bastian, 1965, p. 592).

Lieberman (1984, 1991) has discussed at length the changes in the human vocal tract that have contributed to our extraordinary vocal capabilities. The most important are the descent of the larynx into the neck to create a pharyngeal cavity, the change in the form and position of the tongue, and the right-angle bend between the pharyngeal (throat) and oral (mouth) cavities (Figure 1–1.) The larynx was originally a simple valve to protect the lungs of fish from the influx of water and foreign objects (Negus, 1949, cited in Lieberman, 1984). In ter-

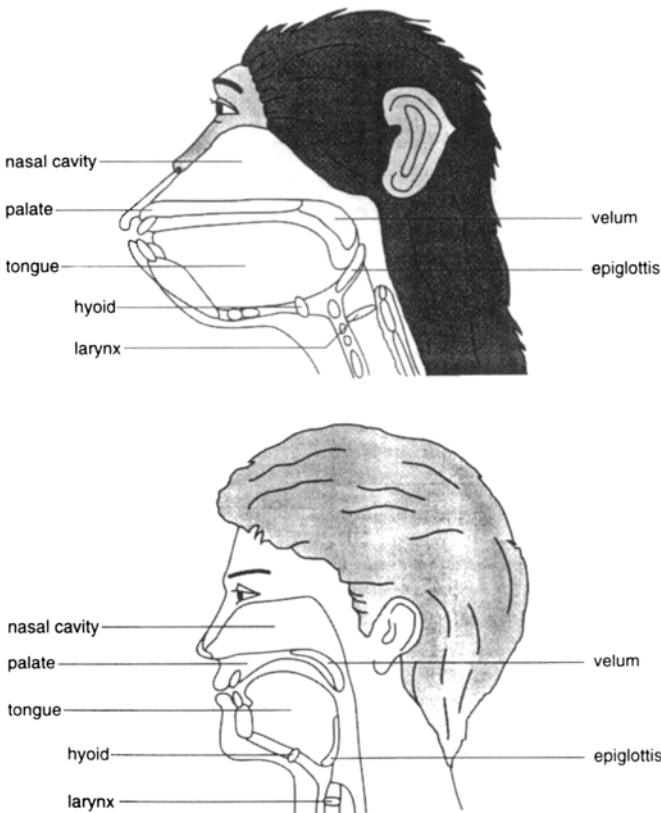


Figure 1–1: Vocal tract of the chimpanzee and human. From Whitney, P. (1998). *The Psychology of Language*. Figure 1.2, p.6. Copyright © 1998 by Houghton Mifflin Company. Used with permission.

restrial mammals, the larynx converts the flow of air from the lungs into phonation. The rate of opening and closing of the vocal cords sets the fundamental frequency of phonation (the lowest pitch) (Lieberman, 1975). The supralaryngeal vocal tract acts as a filter, maximizing acoustic energy at particular frequencies (formants) depending on its general configuration—that is, its length and shape (Lieberman, 1991). The basic structure of the larynx of the great apes is similar to humans' (Negus, 1949). However, apes also have laryngeal air sacs that may serve as resonators during loud vocalizing to compensate for the lack of a pharynx (Marler & Tenaza, 1977). Apes are capable of changing the filtering properties of the supralaryngeal vocal tract to produce variations in formant frequency, and the range of variation is similar to that found in the human neonate (Lieberman, 1975).

At birth, the human infant's vocal tract resembles that of lower mammals in that the larynx is high and can rise to seal off the nasopharynx (nasal cavity). This has the survival advantage of protecting the neonate from choking, because food can pass to either side of the raised larynx into the pharynx (Lieberman, 1984). Very young infants are further protected by being programmed to breathe only through their noses, and they can breathe while they eat, a feat that adult humans cannot (or should not) attempt. Human infants are also born with a thin tongue situated entirely in the mouth. By 3 months of age, the human infant's vocal tract has begun to resemble the human adult's in several ways: (1) the palate begins to move back, (2) the larynx has begun to descend in the neck and to become more mobile, (3) the tongue has become rounder and protrudes into the throat to form a movable anterior wall, and (4) the oral and pharyngeal cavities have become positioned more at right angles. The right-angle bend now helps the velum at the back of the throat to close off the nasal cavity, making possible the more easily perceptible non-nasal sounds (Lieberman, 1991). It may also be important in producing "stop" consonants (those with full occlusion) (Kimura, 1993) and the extreme high vowels /i/ as in *beet* and /u/ as in *toot* (Lieberman, 1975). In addition, the infant's ribs have moved from a perpendicular orientation to the spine to a downward slant, allowing the intercostal muscles and diaphragm to inflate the lungs. This last adaptation is not specifically human, because adult chimpanzees and gorillas also have downward-slanting ribs (Lieberman, 1984).

We know from computer simulations (Lieberman, Klatt, & Wilson, 1969) that the nonhuman primate vocal tract (and the human infant's vocal tract before 3 months) will not allow production of the full range

Cambridge University Press

978-0-521-59299-4 - Routes to Child Language: Evolutionary and Developmental Precursors

Joanna Blake

Excerpt

[More information](#)*Prelinguistic Vocalizations*

9

of vowels, particularly the extremes of the high front vowel /i/, the high back vowel /u/, and the low midvowel /a/ as in *mama*. However, non-human primates do not fully utilize the vowel space that they do have, whereas 3-month-old human infants do (Lieberman, 1984). Thus, the evolution of human speech-producing capacity cannot be entirely explained by changes in the vocal tract. Another important evolutionary change is that nonhuman vocalizations are controlled by the cingulate (old motor) cortex, the basal ganglia, and midbrain structures and not by the neocortex, as they are in humans (Lieberman, 1991, 1995). Neocortical stimulation produces vowel-like sounds in humans but no vocalizations in squirrel or rhesus monkeys, although chimpanzees might be an exception (Kimura, 1993; Ploog, 1988; Ploog & Jurgens, 1980). Most of the vocal repertoire of captive squirrel monkeys can be elicited by stimulating areas within the midbrain and limbic system (Hauser, 1996). We also know from a century of research on brain-damaged patients that Broca's area (Figure 1–2), which may have a structural homologue but no functional counterpart in nonhu-

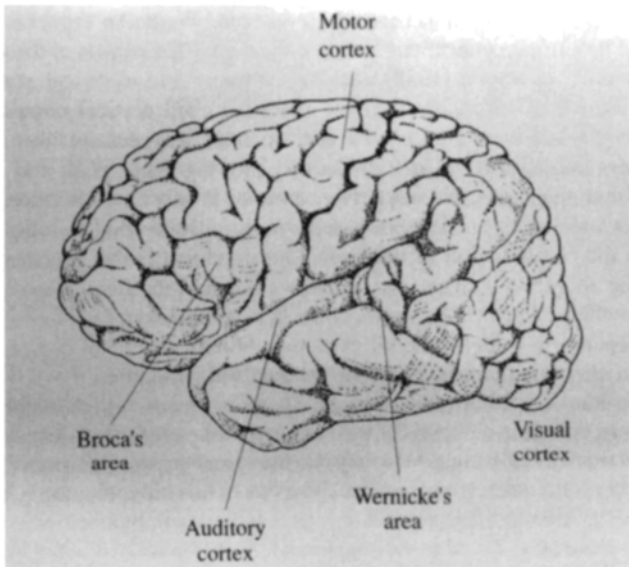


Figure 1–2: Left hemisphere of a human brain showing Broca's and Wernicke's areas. From Noble, W., and Davidson, I. (1996). *Human evolution, language and mind: A psychological and archaeological enquiry* (p. 17). Cambridge: Cambridge University Press. Reprinted with permission from Cambridge University Press.

man primates, is involved in the motor programs necessary for speech production (but see chapter 8). The prefrontal cortex also is clearly implicated in both language and thought (Bates, 1996; Hauser, 1996; Lieberman, 1991, 1995).

Voluntary neocortical control of complex vocal motor patterns thus seems to be restricted to humans. Cortical control means a “measure of independence of the signal from the limbic system and the direct, unconscious expression of individual emotion” (Lancaster, 1968, p. 453). Goodall (1986) has also stressed that chimpanzee calls are closely tied to underlying emotions and that these primates are unable to vocalize at will, even having difficulty in suppressing their vocalizations in situations of danger that require silence. This last statement, however, seems to be contradicted by her observation that chimpanzees on patrol of their territory are quite silent. Learning when to release or inhibit calls appears to be regulated by the anterior cingulate cortex, and rhesus monkeys with lesions to this area are no longer able to master operant conditioning tasks (Jurgens, 1995). Several species of non-human primates have been successfully operantly conditioned to produce a vocalization under certain conditions and not produce it when these conditions are absent in order to receive a reward (Pierce, 1985). Although it is claimed that these findings indicate that nonhuman primate vocalizations are modifiable, the acoustic structure of their vocalizations was not apparently modified, but only their frequency or duration. Lieberman (1995) has proposed that “the ability of humans to ‘free up’ the stereotyped motor acts that make up non-human primate calls derives from prefrontal cortex ‘overriding’ basal ganglia coded patterns” (p. 278). Researchers studying the semioticity of primate calls do not, of course, agree that primate vocalizations are strictly emotive and involuntary. Others argue that a strict dichotomy that posits limbic and involuntary against neocortical and voluntary is not useful (Steklis, 1985). The question of semioticity, and its potential voluntary aspect, is revisited in chapter 2.

It is nevertheless clear that the vocal tract of nonhuman primates does allow them to produce humanlike sounds that they do *not* produce. They *could* occlude their supralaryngeal vocal tracts to produce some of the stop consonants, namely /b/, /p/, /d/, and /t/, and they are capable of producing /s/ as well (Lieberman, 1991). In fact, Hauser and Marler (1992) found that rhesus macaques do use their lips to produce variants of clear tonal calls (“coos”) and tongue movements to produce acoustic changes in alarm barks. These are the articula-