

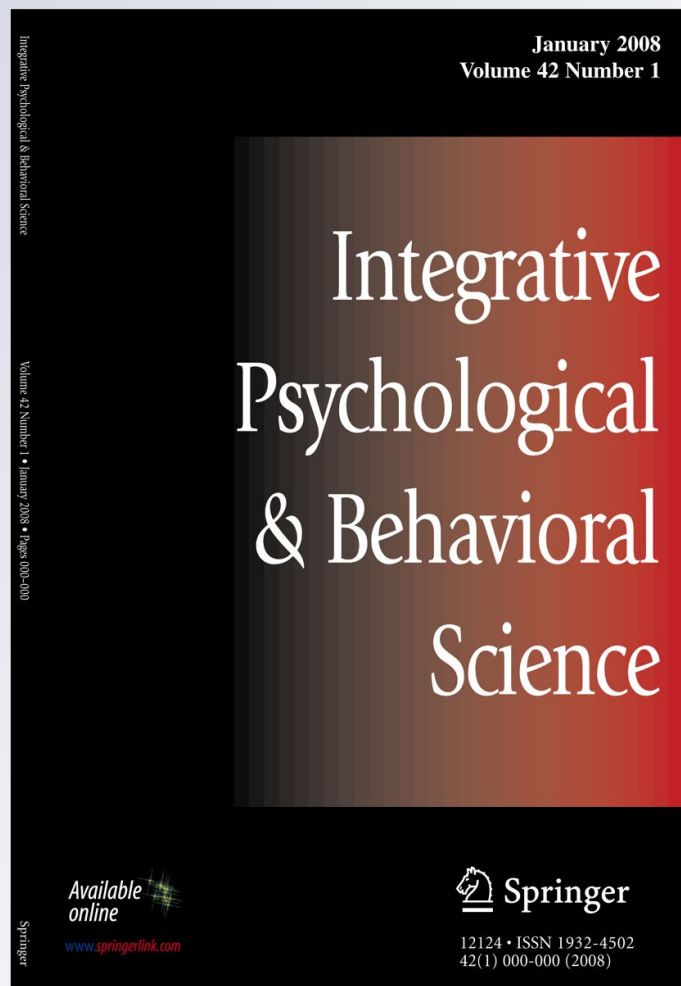
*In the Beginning Was the Familiar
Voice: Personally Familiar Voices in the
Evolutionary and Contemporary Biology
of Communication*

Diana Sidtis & Jody Kreiman

**Integrative Psychological and
Behavioral Science**

ISSN 1932-4502

Integr. psych. behav.
DOI 10.1007/
s12124-011-9177-4



Your article is protected by copyright and all rights are held exclusively by Springer Science+Business Media, LLC. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your work, please use the accepted author's version for posting to your own website or your institution's repository. You may further deposit the accepted author's version on a funder's repository at a funder's request, provided it is not made publicly available until 12 months after publication.

In the Beginning Was the Familiar Voice: Personally Familiar Voices in the Evolutionary and Contemporary Biology of Communication

Diana Sidtis · Jody Kreiman

© Springer Science+Business Media, LLC 2011

Abstract The human voice is described in dialogic linguistics as an embodiment of self in a social context, contributing to expression, perception and mutual exchange of self, consciousness, inner life, and personhood. While these approaches are subjective and arise from phenomenological perspectives, scientific facts about personal vocal identity, and its role in biological development, support these views. It is our purpose to review studies of the biology of personal vocal identity—the familiar voice pattern—as providing an empirical foundation for the view that the human voice is an embodiment of self in the social context. Recent developments in the biology and evolution of communication are concordant with these notions, revealing that familiar voice recognition (also known as vocal identity recognition or individual vocal recognition) has contributed to survival in the earliest vocalizing species. Contemporary ethology documents the crucial role of familiar voices across animal species in signaling and perceiving internal states and personal identities. Neuropsychological studies of voice reveal multimodal cerebral associations arising across brain structures involved in memory, emotion, attention, and arousal in vocal perception and production, such that the voice represents the whole person. Although its roots are in evolutionary biology, human competence for processing layered social and personal meanings in the voice, as well as personal identity in a large repertory of familiar voice patterns, has achieved an immense sophistication.

D. Sidtis
Department of Communicative Sciences and Disorders, New York University, New York, NY 10012,
USA
e-mail: diana.sidtis@nyu.edu

D. Sidtis (✉)
Division of Geriatrics, Nathan Kline Institute for Psychiatric Research, 140 Old Orangeburg Road,
Orangeburg, NY 10962, USA
e-mail: diana.sidtis@nyu.edu

J. Kreiman
Division of Head/Neck Surgery, UCLA School of Medicine, 31–24 Rehab Center, Los Angeles, CA
90095–1794, USA
e-mail: jkreiman@ucla.edu

Keywords Personal voice recognition · Self and voice · Neuropsychology of voice · Evolutionary biology

Voices Are All Around Us: Familiar Voices are Special

Voices are everywhere in human society. Every viable speaker in the world produces a signature voice pattern, and listeners naturally derive an abundance of judgments from it. All kinds of consequences, benign as well as dire, flow from human interactions with voice quality. Understanding these judgments and analyzing these behaviors have interested researchers in an appropriately large array of disciplines, including psychology, linguistics, speech and hearing sciences, forensics, communicative disorders, evolutionary biology, and anthropology. This is true because voices exist and operate in personal, sociological, and historical dimensions. Although these dimensions are manifest only in the speaker-listener interaction, in traditional approaches to the study of voice, speakers and listeners are examined as if they were independent entities linked only by a physical (acoustic) signal. Yet the voice pattern is best approached as constituting a dialogic process. As such, it has evolved to play a major part in the biology of communication, signaling reproductive fitness, fostering mother/infant reunifications and bonding (and hence infant survival), facilitating identification of friend and foe, and enabling the formation of social groups (Locke 2008, 2009; MacLean 1990). Studies of this theme often utilize the term “vocal recognition of identity” (e.g., Rendall et al. 1996). Here the term “familiar voice pattern” is intended to emphasize the roles and properties of this important feature of biology. It is our purpose to describe how it is that the familiar voice—known and recognizable to the listener—has most prominently played these crucial roles in the biology of communication.

This article provides biological foundations to the characterization of voice as embodied social form representing a person and tied to the mutual sharing of self and consciousness (e.g., Bertau 2007, 2008; Osatuke et al. 2004; Linell 2007; Hermans 1996, 1998; Shotter 1996). Related to this view is the elaboration of the multilayered, polyphonic nature of the personal voice pattern (Stiles 1999; Stiles et al. 2004) which has its finest flowering, in interspeaker dialogue, when the familiarity aspect is present. These conceptualizations of voice quality—physical embodiment and multilayering as revealed in the dialogic process of mutually sharing, speaking and listening subjects—are supported by evolutionary and biological perspectives (Kreiman and Sidtis 2011). Evidence is provided here, first, that familiar voice patterns are special in human affairs; that their salient role in infant survival begins even before birth; that inherent in each is an elaborate constellation of biographical information; and that it takes the whole brain and, by extension, the whole person to participate in producing and perceiving a voice. Biological perspectives reveal that synergistic production and perception of familiar voices in the environment have crucially guided survival and development across the millennia of evolution. These mutually shared vocal exchanges arise from the specific cultural contingencies of social groupings (Bakhtin 1973, 1986; Neisser and Jopling 1997; Josephs 2002).

Voice quality processing in humans is a prodigious cognitive ability, second only to language in depth, complexity, and extent. Processing of the voice pattern evolved over millennia of development in many species, but the elaboration of the role of

voice in humans is immense, just as human language has also attained a certain unique complexity. Our research leads to the proposal that it is the personally familiar voice that emerges as crucial to biological, evolutionary and social development. No known limit has been demonstrated for the repertory of recognizable voices in humans. Acquaintances expect to be recognized by voice (Schegloff 1979). It is likely that other vocal information—intentions, emotions, inferences, mood, attitudes—is more competently transmitted when familiar voices are in play (Nygaard 2005). Production and perception of personally familiar voice patterns, representing the multifaceted self in significant behavioral and cultural contexts, antedated by a very long time the development of speech and language, and allowed for social relating between individuals in earliest evolutionary times. Better understanding of the role of familiarity in human neuropsychology has come only recently, because utilizing familiar stimuli in controlled studies poses special challenges (Van Lancker 1991; Kreiman and Sidtis 2011).

One might suppose that recognizing a familiar voice arises from a more fundamental ability to discriminate among unfamiliar voices (e.g., Fischer 2004). However, convergence of several observations suggests that familiar voice recognition is the more elemental process. First, persons with focal brain damage are able to recognize familiar voices despite a neuropsychological impairment in the ability to discriminate among unfamiliar voices (Neuner and Schweinberger 2000; Van Lancker and Kreiman 1986, 1987; Van Lancker et al. 1988). The reverse is also true, suggesting that in the adult, the two abilities may exist as independent and unordered in processing, and that, although both abilities are based in fundamental features of auditory-acoustic processing, recognition does not necessarily depend on or arise from discrimination. Interactions of the two functions at several stages of auditory processing and during acquisition and development, of course, are to be expected.

Secondly, as further evidence of the primacy of the familiar voice, the ability to recognize the voice of one's mother is present at birth in normally hearing humans (DeCasper and Fifer 1980; Querleu et al. 1984; Hepper et al. 1993; Mehler et al. 1978), in contrast to the ability to discriminate among unfamiliar voices, which does not fully develop until about age 10 (Mann et al. 1979; Mehler et al. 1978). The role of mother-infant mutually instantiated vocal recognition in establishing personal ingredients as building blocks of self, social development and acculturation can hardly be exaggerated (Bertau 2007). Neurophysiological evidence confirms that the mother's voice has special meaning to an infant, in that mother's voice excites unique brain patterns (Purhonen et al. 2004, 2005). Recognizing and responding to the mother's voice is an important ingredient for fostering infant/mother bonding, which in turn is essential for the survival of the infant given its total dependence on its mother (Locke and Bogin 2006). In fact, recognition of mothers by infants, infants by mothers, and mutual voice recognition between parents and offspring are all common in animals, as discussed below ([In the Beginning Was the Voice Pattern: Voice Recognition Evolved Early and Across Many Species](#) section).

Beyond early infancy, familiar voices retain their special status as signifiers of multilayered information. The personally familiar voice as an emotional auditory event travels via specialized auditory pathways directly to subcortical mechanisms, which, according to the Cannon-Bard theory, arouse corporeal and cerebral representations of emotion (Damasio 1994). The voice carries and evokes

continuous flickers of emotion and attitude, which arise from associated evaluations stored in memory that “become active automatically on the mere presence or mention of the object in the environment” (Bargh et al. 1992, p. 893). Implicit memorial associations of dizzying variety attach to the familiar voice pattern (Fig. 1). In contrast, inner experiencing of an unfamiliar voice is relatively impoverished. The subjective emotions associated with familiar voices, together with the prosodic qualities carried in the voice signal at the time of hearing the voice, combine to form a powerful vehicle to acquire, process, and recall the personally familiar voice pattern (Laird et al. 1982). (We have elsewhere described the rapid acquisition of the familiar voice pattern, due to engagement of arousal and emotional systems in the brain (Kreiman and Sidtis 2011, Chapter 6).) These ingredients contribute to the expression of self and personhood in the voice and to the mutual sharing of self in communicative interaction.

It Takes a Whole Brain to Produce and Recognize a Voice

The whole self underlies the personal vocal pattern, such that a large range of subtle, ephemeral personal characteristics are encoded and conveyed to the listener, who infers these attributes from voice. This includes the physical self (gender and age, health, reproductive fitness, race, size, and attractiveness; see Kreiman and Sidtis 2011, chapter 4), the social self (education, background), and the speaker’s personhood (personality, mood, emotions, and attitudes) (Laver 1968; Berry 1990; Revelle and Scherer 2009; Scherer 1986; Konopcznski 2010). Producing and perceiving these characteristics relies on disparate brain structures that modulate

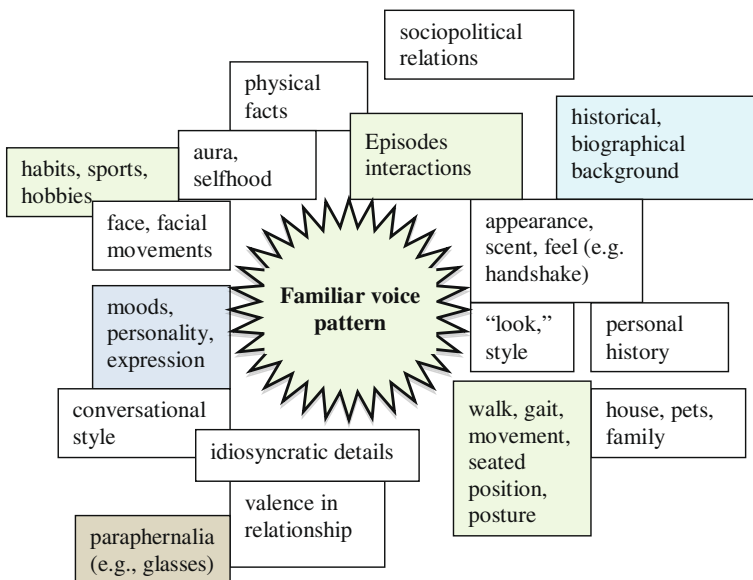


Fig. 1 A subjective, unstructured array of qualities accompanies the known voice pattern as personally relevant factors

cross-modal associations, memory, attention, arousal and emotion (Andersen 1997; Fair 1988, 1992; Kayser and Logothetis 2007; Schroeder, et al. 2003). These facts provide the foundation for the notion that the voice is revelatory of “self,” mental states, and consciousness, and reflects both the speaker and the context in which the voice is produced. The association of voice with self, consciousness, and in earlier times, with soul, enjoys a venerable tradition since at least the time of Aristotle (von Kempelen 1791; Rush 1823; Kidd 1857; Konopcznski 2010) and is enfolded in the goals of contemporary voice coaching and therapy (e.g., Linklater 1976; Boone 1991). Modern perspectives describe how vocal signals reflect the listener’s assumptions, expectations, experiences with the speaker, and cognitive perspectives (Pollermann 2010). In this conceptualization, it is largely through use of the voice that the self of each co-participant is mutually shared in communicative interaction (Bertau 2008); the embodied voice arises from the whole person. All dimensions of the voice enter into this process for the listener, but little neuroscientific research has focused on the manner in which voice arises from such complex social and psychological processes, so that much of the communication that passes between interlocutors via the voice remains undescribed and unexplained. In this section, we provide grounding for these ideas derived from recent developments in cerebral function.

In describing the process of vocal production, studies of voice usually focus on a few cerebral structures: the motor strip of the cortex, selected motor pathways through the basal ganglia, and a series of cranial nerves that activate laryngeal and vocal tract muscles. It is more accurate to say that producing a voice in humans involves, in addition to these structures, the midbrain, most subcortical nuclei and limbic structures, the cortical lobes and the cerebellum (Kreiman and Sidtis 2011; Jürgens 2002). One source of this proposition arises directly from the effects on voice of neurological damage or experimental stimulation: the periaqueductal gray matter in the midbrain (Esposito et al. 1999), several basal ganglia and limbic (subcortical) structures (Simonyan and Jürgens 2003; Damasio 1994), temporal, parietal and frontal lobes, and the cerebellum are all involved in normal voice production (Ackermann et al. 2007). Indirect evidence lies in the respective neuropsychological roles of these structures: initiating behaviors, modulating motor gestures and emotion, monitoring auditory input, engaging pertinent cognitive associations (Fair 1992), and managing ongoing action (Masterman and Cummings 1997). Similarly for voice perception, traditional neuroanatomical descriptions depict the auditory pathway as a through-put channel and the auditory receiving areas of the cortex as providing interpretation of sound. In fact, perceiving a voice engages many levels of processing all along the auditory pathway, and visual, somatosensory and auditory signals play key roles at early, low-level stages of auditory cortical processing (Schroeder et al. 2003). This confluence of auditory and multisensory streams precedes cognitive processing of sound (Winer and Lee 2007). It is now recognized that sound processing in the auditory cortex occurs not in one receiving center, but in various cortical fields, and further, that these areas interact richly with other areas (Petkov et al. 2006; Winer and Lee 2007) through numerous connections with subcortical nuclei and cross model association areas of the frontal and parietal lobes (Fair 1992; Benson 1994; Haramati et al. 2008). This picture of interactional, whole brain processing of sound is compatible with current models of the voice as infused with affect, thought, memory, motivation and attention, via a

web of reciprocal influences involving multisensory areas and nuclei along with extensive cerebral connections (Miall 1986). Thus the voice pattern arises from coordinated systems in the whole brain, which manifest the collective perspectives, knowledge domains and experiences of the speaker, deriving from an integrated network of limbic, cortical, subcortical, cerebellar and brainstem structures that underlie and modulate vocal production and perception and contribute to each produced or perceived voice pattern.

While neuropsychological research was focused for several decades on the contributions of the cortex to human behavior, there is currently a resurgence of interest in limbic structures, the seat of emotion, with neuroscientists pointing out that all of human behavior is richly imbued with emotional tone (e.g., Panksepp 1998, 2003). Interest in subcortical structures, which are now known to modulate motivation, attention, and initiation of motor behaviors, has also grown (Marsden 1982; Bhatia and Marsden 1994; Saint-Cyr et al. 1995; Lieberman 2002). This perspective is concordant with current descriptions in neuroscience of “large-scale neurocognitive networks” contributing to mental activity that is highly flexible and almost infinitely rich (Mesulam 1990, p. 597), and with a movement away from localization of functional representation to descriptions of functional networks in the brain (Nudo et al. 2001). This modern representation of voice production and perception is consonant with dialogic conceptualizations of voice as reflective of self in the broadest sense and with multilayering of personal information (Osatuke et al. 2005; Bertau 2007; Hermans 1998).

In the Beginning Was the Voice Pattern: Voice Recognition Evolved Early and Across Many Species

Ethological and evolutionary data indicate that, in addition to use of various call repertoires for semantic signaling purposes, recognition of familiar voice patterns, which include crucial information about the physical size, gender, emotions, mood, and subtleties of intention, is a widespread ability that has evolved in response to different environmental and behavioral demands. Monkeys and primates, bats, penguins, sheep, goats, deer, horses, wolves, frogs, elephants, and birds recognize the familiar voices of parent and/or child, or of a neighbor. Based on its modern distribution, this ability had evolved by the time that frogs appeared and before the advent of mammals (Burke and Murphy 2007; Bee et al. 2001). The males of a variety of anuran species (including the North American bullfrog, aromabatid frogs, and concave-eared torrent frogs; Bee and Gerhardt 2002; Gasser et al. 2009; Feng et al. 2009) are able to discriminate the voices of familiar neighbors from those of unfamiliar frogs, and use this information to determine the level of aggressive response necessary to defend their mating grounds. Males respond significantly less aggressively to a familiar voice than to an unfamiliar voice, independent of the location from which the call sounds, suggesting that they can discriminate the familiar voices of their neighbors from unfamiliar voices, and can use this information to limit unnecessary aggressive interactions (Bee and Gerhardt 2002).

Producing and recognizing familiar voice patterns thus antedates by millions of years the other more celebrated evolutionary developments in communication and

cognition. Although voice recognition abilities are shared by many animals, the manner in which animals recognize one another varies widely in form and complexity from species to species. In recent times, vocal identity studies, or recognition of the individual by voice, have proliferated in the scientific literature (see Kreiman and Sidtis 2011). In some species, voice is the primary means by which individuals recognize each other, as in penguins (Jouventin 1982); or it may be used as an adjunct to visual and/or olfactory cues, as in sheep (Searby and Jouventin 2003) and goats (Terrazas et al. 2003). Recognition may be mutual, so that infants and parents recognize each other, or asymmetrical, so that parents recognize offspring or offspring parents, but not both. For example, fallow deer fawns hide from predators, and thus must remain silent until summoned by their mothers; thus, fawns can recognize the calls of their own mothers but mothers do not recognize fawns (Torriani et al. 2006). In contrast, ewes and their lambs, who live in herds, can each recognize the other based solely on their calls (Searby and Jouventin 2003). Finally, calls can be relatively simple in structure in animals that breed in small herds, follow their parents (as sheep and goats do), or rely on additional cues to recognize family members. For example, the exact role that voice plays in recognition varies across penguin species, and call complexity varies accordingly (Jouventin and Aubin 2002; Searby et al. 2004). In nesting (Adélie, macaroni, and gentoo) penguins, the nest location is the primary cue to the penguin's identity, so calls serve mostly to summon a chick to the nest when a parent returns from foraging and are accordingly quite simple in acoustic structure, with recognition depending primarily on pitch (Jouventin and Aubin 2002). In contrast, emperor and king penguins incubate their eggs on their feet, so recognition depends completely on voice quality. Accordingly, their calls are acoustically highly complex (Aubin et al. 2000; Lengagne et al. 2001).

Various types of voice patterns and recognition strategies have evolved in other species that breed in colonies, produce mobile offspring, and/or forage, and thus must rely on voice recognition if offspring are to survive (e.g., Searby et al. 2004). For example, seals breed in colonies of up to 70,000 animals, and because feeding grounds are often far from breeding shores, mothers must leave pups unattended sometimes for weeks while foraging for food. Mothers recognize pups by voice in seven different species of seal; in four of these, pups also recognize their mothers (Insley 2001). While pups' voices change as they grow, mother seals retain the ability to identify these different versions produced at different ages of development (Charrier et al. 2001, 2003). Voice patterns may also be inborn: Evening bat mothers go out to forage immediately after the birth of their young, but successfully reunite with their pups in dark crèches that may contain thousands of other bats, suggesting calls have a genetic component (Scherrer and Wilkinson 1993). Although unilateral or mutual voice recognition is usually instantaneous for offspring, learning to recognize a parent's or infant's voice may follow a delayed course in animals whose environment and/or behavior makes voice recognition less critically important. In some primate species the responsibility for reunification rests on mothers, and infants may not recognize their parent until well after birth (Altmann 1980). Japanese macaques do not recognize their mothers until about 22 days of age (Masataka 1985), and infant barbary macaques apparently do not recognize their mothers reliably until 10 weeks of age (Fischer 2004). Although studies in the field confirm familiar voice pattern recognition behaviors in nonhuman primates (Cheney

and Seyfarth 1980, 1999; Rendall et al. 1996), the repertory of recognized identities in other species studied up to this time is usually limited to a very few at a time (Hansen 1976). One exception lies in elephant society, where many individuals in surrounding herds—up to a hundred—are recognized by voice (McComb et al. 2002). Further studies may reveal more about this ability in nonhuman animals. In the meantime, human management of exquisitely layered prosodic meanings within an immense expanse of personal identities, signaled by the voice pattern, remains an astonishing and dazzling biological tour de force.

In summary, the wide-spread presence and sophistication of familiar voice recognition across species, along with the variety of recognition behaviors that has evolved, attest to the fundamental importance of this ability in biology (MacLean 1990). The details of the observed variations can often be understood in terms of the varying ecological and evolutionary demands faced by different animal species. These facts underscore the importance of the ability to recognize a familiar voice, the essential mutuality of the process, and the special status that familiar voices have in biology and animal behavior.

A Neuropsychological-Social Model of Voice Recognition

Acquiring and recognizing a familiar voice involves tapping into an extensive network of attributes associated with the target voice, and draws on a constellation of diverse characteristics, including affective and attitudinal qualities, biographical history, appearance, dress, gait, unique markings and customary paraphernalia, preferences, and so on, which are stored together in a broad based, integral Gestalt percept (see Fig. 1). Subjective qualities in philosophical treatments of consciousness, self, and awareness (Buck 1993) emerge most prominently for the familiar voice pattern, as essentially different from the status and role of unfamiliar voices. A final common pathway for vocal pattern recognition occurs in the right hemisphere, site of personal relevance (Van Lancker and Canter 1982; Van Lancker and Kreiman 1986; Belin et al. 2000; von Kriegstein and Giraud 2004; Sidtis and Kreiman 2008). In contrast, discriminating unfamiliar voices utilizes feature-analysis and feature-matching to a general template or set of templates that approximately fit a set of perceived auditory features, and is more successfully modulated by the left cerebral hemisphere and/or by both hemispheres (Kreiman 1997; Kreiman and Sidtis 2011). This characterization of voice perception is concordant with known differences in cerebral processing emerging from clinical and experimental data, with familiarity and pattern recognition represented in the right hemisphere, and temporal and analytic processes more successfully processed in the left cerebral hemisphere (Bever 1975; Bradshaw and Mattingly 1995; Gazzaniga et al. 2002; Van Lancker 1997).

Loosely adapting psychological approaches that contrast exemplar and rule-based models of learning and memory, we describe these differences with a “Fox and Hedgehog” model of voice recognition, named after the parable in which the fox knows many little things, while the hedgehog knows one big thing (Berlin 1953/1994). This model proposes that perceptual-auditory features (lots of little things) are utilized predominantly for perceiving unfamiliar voices, while for recognition of a familiar

voice, a few signature features suffice to herald the complete, known voice pattern (one big thing) (Kreiman and Sidtis 2011, Chapter 6).

This view of vocal quality perception is consistent with microgenetic or process theory (Brown 1998a, b), which provides a distinctive counterpart to paradigms of brain function based on assumptions about localization of neuropsychological abilities. Microgenesis refers to psychodynamic processes unfolding in a present-time scale from global to local instantiation of a realized percept or experienced mental event (Werner 1956; Rosenthal 2004). A key feature is that form, meaning and value are not independent, but unfold simultaneously from resources within the entire brain (Brown 1988). Our perspective on familiar voices is resonant with the process model of brain function in which configurations play a major role as original status of the cognitive content (Benowitz et al. 1990; Brown 1998b). The mental content is “not constructed like a building,” but unfolds from “preliminary configurations [that] are implicit in the final object” (Brown 2002, p. 8). Affect and familiarity, essential characteristics of familiar voice patterns, are better accommodated in these approaches (Brown 1998b, 2002), which lend themselves to a discussion of affective, subjective, and personally familiar phenomena (Van Lancker 1991). This approach provides a vehicle for characterizing the intimate listener-speaker dyad inherent in familiar voice processing, and for the fact that the voice is expressive of the entire self.

Familiar Voice Patterns Enable Relationships in Sentient¹ Life

Voice recognition competence, or successful recognition of vocal identity, is present in many species and present at birth in humans, and as an evolutionarily “old” competence takes its place among the most primordial of behaviors. Following Gibson’s (1966) view that sensitivity to biologically useful information evolves with the organism, it follows that sensitivities for acquiring a familiar voice pattern arise from the most basic perceptual processes.

The pervasiveness and importance of the human voice in the psychology and sociology of life, and the breadth of its effects on human behavior, cannot be overstated, and the special status of the familiar voice pattern has been described here.² There are seldom events in daily living that do not involve production and perception of voice, with the attendant conscious and unconscious judgments of the talker by the listener. Personally relevant voices, by definition, are represented in memory with emotional reference to the self. Subjective impressions of voice as embodiment of self are strongly supported by the biological facts presented here. The intense personal meanings of voice have a profoundly important role in animal behavior, and continue to weave subtle strands of communication everywhere in modern human life. Social relating across biological species has relied on familiar voice patterns since the very beginnings of vocal behavior. Friend or foe, offspring and parents, family and cohort members can be identified at night or at a distance not

¹ The notion of sentience refers here to some form of physiological state of existence and capacity for thinking and/or feeling.

² Many of the characteristics of voices hold also for faces (see Chapter 6, Kreiman and Sidtis 2011).

by visual inspection, but by the voice pattern. The meaning and use of the vocal pattern grew with the growth in the complexity of cognition and social organization of various species. The prodigious presence of familiar voice patterns in human experience, and its role in facilitating mutual exchange of the inner voice, has its foundations in an evolutionary past shared by many other animals. However, the evolutionary leap from observable animal behaviors to the present state of human competence in voice recognition is as great as that from animal calls to human language. In humans, as described by Bertau (2008), the evolution of the personally relevant voice pattern accompanies and enhances the development of consciousness and self-awareness, as well as empathy for and recognition of the other.

Acknowledgements Preparation of this paper was supported in part by grant DC01797 and by R01 DC007658 from the National Institute on Deafness and Other Communication Disorders.

References

- Ackermann, H., Mathiak, K., & Riecker, A. (2007). The contribution of the cerebellum to speech production and speech perception, clinical and functional imaging data. *Cerebellum*, 6, 202–213.
- Altmann, J. (1980). *Baboon mothers and infants*. Chicago: University of Chicago Press.
- Andersen, R. A. (1997). Multimodal integration for the representation of space in the posterior parietal cortex. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 352, 1421–1428.
- Aubin, T., Jouventin, P., & Hildebrand, C. (2000). Penguins use the two-voice system to recognize each other. *Proceedings of the Royal Society of London B, Biological Sciences*, 267, 1081–1087.
- Bakhtin, M. (1973). *Problems of Dostoevsky's poetics* (2nd ed.), translated by R. W. Rotsel. Ann Arbor: Ardis.
- Bakhtin, M. M. (1986). *Speech genres and other late essays*, translated by V. W. McGee. Austin: University of Texas Press.
- Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality and Social Psychology*, 62, 893–912.
- Bee, M. A., & Gerhardt, H. C. (2002). Individual voice recognition in a territorial frog (*Rana catesbeiana*). *Proceedings of the Royal Society of London B, Biological Sciences*, 269, 1443–1448.
- Bee, M. A., Kozich, C. E., Blackwell, K. J., & Gerhardt, H. C. (2001). Individual variation in advertisement calls of territorial male green frogs, *Rana clamitans*, Implications for individual discrimination. *Ethology*, 107, 65–84.
- Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, 403, 309–312.
- Benowitz, L. I., Finkelstein, S., Levine, D. N., & Moya, K. (1990). The role of the right cerebral hemisphere in evaluating configurations. In C. B. Trevarthen (Ed.), *Brain circuits and functions of the mind* (pp. 320–333). Cambridge: Cambridge University Press.
- Benson, D. F. (1994). *The neurology of thinking*. Oxford: Oxford Univ. Press.
- Berlin, I. (1953/1994). *The hedgehog and the fox, an essay on Tolstoy's view of history*. London: Weidenfeld and Nicolson. (Reprinted in *Russian Thinkers*, Oxford: Penguin).
- Berry, D. S. (1990). Vocal attractiveness and vocal babyishness: Effects on stranger, self, and friend impressions. *Journal of Nonverbal Behavior*, 14, 141–153.
- Bertau, M.-C. (2007). On the notion of voice, An exploration from a psycholinguistic perspective with developmental implications. *International Journal for Dialogical Science*, 2, 133–161.
- Bertau, M.-C. (2008). Voice, A pathway to consciousness as 'social contact to oneself'. *Integrative Psychological and Behavioral Science*, 42, 92–113.
- Bever, T. G. (1975). Cerebral asymmetries in humans are due to the differentiation of two incompatible processes, Holistic and analytic. *Annals of the New York Academy of Science*, 263, 251–262.
- Bhatia, K. P., & Marsden, C. D. (1994). The behavioural and motor consequences of focal lesions of the basal ganglia in man. *Brain*, 117, 859–876.
- Boone, D. (1991). *Is your voice telling on you?* San Diego: Singular.
- Bradshaw, J. L., & Mattingly, J. B. (1995). *Clinical neuropsychology, behavioral and brain science*. New York: Academic.

- Brown, J. W. (1988). *The life of the mind, selected papers*. Hillsdale, NJ: Lawrence Erlbaum.
- Brown, J. W. (1998a). Morphogenesis and mental process. In C. Pribram & J. King (Eds.), *Learning as self organization* (pp. 295–310). Mahwah: Lawrence Erlbaum.
- Brown, J. W. (1998b). Foundations of cognitive metaphysics. *Process Studies*, 21, 79–92.
- Brown, J. W. (2002). *The self embodying mind*. Barrytown: Station Hill Press.
- Buck, R. (1993). What is this thing called subjective experience? Reflections on the neuropsychology of qualia. *Neuropsychology*, 7, 490–499.
- Burke, E. J., & Murphy, C. G. (2007). How female barking tree frogs, *Hyla gratiosa*, use multiple call characteristics to select a mate. *Animal Behaviour*, 74, 1463–1472.
- Charrier, I., Mathevon, N., & Jouventin, P. (2001). Mother's voice recognition by seal pups. *Nature*, 412, 873.
- Charrier, I., Mathevon, N., & Jouventin, P. (2003). Individuality in the voice of fur seal females, An analysis study of the pup attraction call in *Arctocephalus tropicalis*. *Marine Mammal Science*, 19, 161–172.
- Cheney, D. L., & Seyfarth, R. (1980). Vocal recognition in free-ranging vervet monkeys. *Animal Behaviour*, 28, 362–367.
- Cheney, D. L., & Seyfarth, R. M. (1999). Recognition of other individuals' social relationships by female baboons. *Animal Behaviour*, 58, 67–75.
- Damasio, A. R. (1994). *Descartes' error*. New York: Avon Books.
- DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: newborns prefer their mothers' voice. *Science*, 208, 1174–1176.
- Esposito, A., Demeurisse, G., Alberti, B., & Fabbro, F. (1999). Complete mutism after midbrain periaqueductal gray lesion. *NeuroReport*, 10, 681–685.
- Fair, C. M. (1988). *Memory and central nervous system organization*. New York: Paragon House.
- Fair, C. M. (1992). *Cortical memory functions*. Boston: Birkhäuser.
- Feng, A. S., Arch, V. S., Yu, Z., Yu, X.-J., Xu, Z.-M., & Shen, J.-X. (2009). Neighbor–Stranger discrimination in concave-eared Torrent Frogs, *Odorrana tormota*. *Ethology*, 115, 851–856.
- Fischer, J. (2004). Emergence of individual recognition in young macaques. *Animal Behaviour*, 67, 655–661.
- Gasser, H., Amézquita, A., & Hödl, W. (2009). Who is calling? Intraspecific call variation in the aromobatid frog *Allobates femoralis*. *Ethology*, 115, 596–607.
- Gazzaniga, M. S., Irvy, R. B., & Mangun, G. R. (2002). *Cognitive neuroscience, biology of the mind*. New York: Norton & Company.
- Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.
- Hansen, E. W. (1976). Selective responding by recently separated juvenile rhesus monkeys to the calls of their mothers. *Developmental Psychobiology*, 9, 83–88.
- Haramati, S., Soroker, N., Dudai, Y., & Levy, D. A. (2008). The posterior parietal cortex in recognition memory, A neuropsychological study. *Neuropsychologia*, 46, 1756–1766.
- Hepper, P. G., Scott, D., & Shahidullah, S. (1993). Newborn and fetal response to maternal voice. *Journal of Reproductive and Infant Psychology*, 11, 147–153.
- Hermans, H. J. M. (1996). Voicing the self, From information processing to dialogical interchange. *Psychological Bulletin*, 119, 31–50.
- Hermans, H. J. M. (1998). The polyphony of the mind, A multivoiced and dialogical self. In J. Rowan & M. Cooper (Eds.), *The plural self, polypsychic perspectives*. Thousand Oaks, CA: Sage Publications.
- Insley, S. J. (2001). Mother-offspring vocal recognition in northern fur seals is mutual but asymmetrical. *Animal Behaviour*, 61, 129–137.
- Josephs, I. (2002). 'The Hopi in Me'. The construction of a voice in the dialogical self from a cultural psychological perspective. *Theory & Psychology*, 12, 162–173.
- Jouventin, P. (1982). Visual and vocal signals in penguins, their evolution and adaptive characters. *Advances in Ethology*, 24, 1–149.
- Jouventin, P., & Aubin, T. (2002). Acoustic systems are adapted to breeding ecologies, Individual recognition in nesting penguins. *Animal Behaviour*, 64, 747–757.
- Jürgens, U. (2002). Neural pathways underlying vocal control. *Neuroscience and Biobehavioral Reviews*, 26, 235–258.
- Kayser, C., & Logothetis, N. K. (2007). Do early sensory cortices integrate cross-modal information? *Brain Structure & Function*, 212, 121–132.
- Kidd, R. (1857). *Vocal culture and elocution*. Cincinnati, OH: Van Antwerp, Bragg & Co.
- Konopcznski, G. (2010). Les enjeux de la voix. In M. R. Castarede & G. Konopcznski (Eds.), *Au commencement était la voix* (pp. 33–52). Toulouse: Erès.
- Kreiman, J. (1997). Listening to voices: Theory and practice in voice perception research. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 85–108). New York: Academic.

- Kreiman, J., & Sidtis, D. (2011). *Foundations of voice studies: Interdisciplinary approaches to voice production and perception*. Boston: Wiley-Blackwell.
- Laird, J. D., Wagener, J. J., Halal, M., & Szegda, M. (1982). Remembering what you feel, effects of emotion on memory. *Journal of Personality and Social Psychology*, *42*, 646–657.
- Laver, J. (1968). Voice quality and indexical information. *British Journal of Disorders of Communication*, *3*, 43–54.
- Lengagne, T., Lauga, J., & Aubin, T. (2001). Intra-syllabic acoustic signatures used by the king penguin in parent–chick recognition, An experimental approach. *Journal of Experimental Biology*, *204*, 663–672.
- Lieberman, P. (2002). *Human language and our reptilian brain: The subcortical bases of speech, syntax, and thought*. Cambridge, MA: Harvard University Press.
- Linell, P. (2007). On Bertau's and other voices (Commentary on Bertau). *International Journal for Dialogical Science*, *2*, 163–168.
- Linklater, K. (1976). *Freeing the natural voice*. Hollywood, CA: Drama Publishers.
- Locke, J. L. (2008). Cost and complexity, Selection for speech and language. *Journal of Theoretical Biology*, *251*, 640–652.
- Locke, J. L. (2009). Evolutionary developmental linguistics, Naturalization of the faculty of language. *Language Sciences*, *31*, 33–59.
- Locke, J. L., & Bogin, B. (2006). Language and life history, A new perspective on the development and evolution of human language. *The Behavioral and Brain Sciences*, *29*, 259–325.
- MacLean, P. D. (1990). *The triune brain in evolution*. New York: Plenum.
- Mann, V. A., Diamond, R., & Carey, S. (1979). Development of voice recognition, Parallels with face recognition. *Journal of Experimental Child Psychology*, *27*, 153–165.
- Marsden, C. D. (1982). The mysterious motor function of the basal ganglia, The Robert Wartenberg Lecture. *Neurology*, *32*, 514–539.
- Masataka, N. (1985). Development of vocal recognition of mothers in infant Japanese macaques. *Developmental Psychobiology*, *18*, 107–114.
- Masterman, D. L., & Cummings, J. L. (1997). Frontal-subcortical circuits, the anatomic basis of executive, social, and motivated behaviors. *Journal of Psychopharmacology*, *11*, 107–114.
- McComb, K., Moss, C., Sayialel, S., & Baker, L. (2002). Unusually extensive networks of vocal recognition in African elephants. *Animal Behaviour*, *59*, 1103–1109.
- Mehler, J., Bertoncini, J., Barriere, M., & Jassik-Gerschenfeld, D. (1978). Infant recognition of mother's voice. *Perception*, *7*, 491–497.
- Mesulam, M.-M. (1990). Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Annals of Neurology*, *28*, 597–613.
- Miall, D. S. (1986). Emotion and the self: The context of remembering. *British Journal of Psychology*, *77*, 389–397.
- Neisser, U., & Jopling, D. A. (1997). *The conceptual self in context: Culture, experience, self-understanding*. Cambridge: Cambridge University Press.
- Neuner, F., & Schweinberger, S. R. (2000). Neuropsychological impairments in the recognition of faces, voices, and personal names. *Brain and Cognition*, *44*, 342–366.
- Nudo, R., Plautz, E. J., & Frost, S. B. (2001). Role of adaptive plasticity in recovery of function after damage to motor cortex. *Muscle & Nerve*, *24*, 1000–1019.
- Nygaard, L. C. (2005). Linguistic and paralinguistic factors in speech perception. In D. B. Pisoni & R. E. Remez (Eds.), *Handbook of speech perception*. Oxford: Blackwell Publishers.
- Osatuke, K., Gray, M. A., Glick, M., Stiles, W. B., & Barkham, M. (2004). Hearing voices. Methodological issues in measuring internal multiplicity. In H. J. M. Hermans & G. Dimaggio (Eds.), *The dialogical self in psychotherapy* (pp. 237–254). New York: Brunner-Routledge.
- Osatuke, K., Humphreys, C. L., Glick, M., Graff-Reed, R. L., McKenzie Mack, L., & Stiles, W. B. (2005). Vocal manifestations of internal multiplicity, Mary's voices. *Psychology and Psychotherapy, Theory, Research and Practice*, *78*, 21–44.
- Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. New York: Oxford University Press.
- Panksepp, J. (2003). At the interface of affective, behavioral and cognitive neurosciences, Decoding the emotional feelings of the brain. *Brain and Cognition*, *52*, 4–14.
- Petkov, C. I., Kayser, C., Augath, M., & Logothetis, N. K. (2006). Functional imaging reveals numerous fields in the monkey auditory cortex. *PLoS Biology*, *4*, e215.
- Pollermann, B. Z. (2010). Qu'exprime la prosodie affective: l'état du corps ou l'état de l'esprit? Proposition d'un modèle de l'émotion et de cognition. In M. F. Castarede & G. Konopczynski (Eds.), *Au commencement était la voix* (pp. 97–104). Toulouse: Erès.

- Purhonen, M., Kilpelainen-Lees, R., Valkonen-Korhonen, M., Karhu, J., & Lehtonen, J. (2004). Cerebral processing of mother's voice compared to unfamiliar voice in 4-month-old infants. *International Journal of Psychophysiology*, *52*, 257–266.
- Purhonen, M., Kilpelainen-Lees, R., Valkonen-Korhonen, M., Karhu, J., & Lehtonen, J. (2005). Four-month-old infants process own mother's voice faster than unfamiliar voices—Electrical signs of sensitization in infant brain. *Cognitive Brain Research*, *24*, 627–633.
- Querleu, D., Lefebvre, C., Titran, M., Renard, X., Morillion, M., & Crepin, G. (1984). Réactivité du nouveau-né de moins de deux heures de vie à la voix maternelle. *Journal de Gynécologie, Obstétrique et Biologie de la Reproduction*, *13*, 125–134.
- Rendall, D., Rodman, P. S., & Edmond, R. E. (1996). Vocal recognition of individuals and kin in free-ranging rhesus monkeys. *Animal Behaviour*, *51*, 1007–1015.
- Revelle, W., & Scherer, K. (2009). Personality and emotion. In D. Sander & K. Scherer (Eds.), *Oxford companion to emotion and the affective sciences* (pp. 304–306). Oxford: Oxford University Press.
- Rosenthal, V. (2004). Microgenesis, immediate experience and visual processes in reading. In A. Carsetti (Ed.), *Seeing, thinking and knowing*. Dordrecht: Kluwer.
- Rush, J. (1823). *The Philosophy of the Human Voice, embracing its physiological history; together with a system of principles, by which criticism in the art of elocution may be rendered intelligible and instruction, definite and comprehension to which is added a brief analysis of song and recitative* (5th ed.). Philadelphia: J.B. Lippencott & Co.
- Saint-Cyr, J. A., Taylor, A. E., & Nicholson, K. (1995). Behavior and the basal ganglia. In W. J. Weiner & A. E. Lang (Eds.), *Behavioral neurology of movement disorders* (pp. 1–28). New York: Raven Press.
- Schegloff, E. A. (1979). Identification and recognition in telephone conversation openings. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 23–78). New York: Irvington.
- Scherer, K. R. (1986). Vocal affect expression, A review and a model for future research. *Psychological Bulletin*, *99*, 143–165.
- Scherrer, J. A., & Wilkinson, G. S. (1993). Evening bat isolation calls provide evidence for heritable signatures. *Animal Behaviour*, *46*, 847–860.
- Schroeder, C. E., Smiley, J., Fu, K. G., O'Connell, M. N., McGinnis, T., & Hackett, T. A. (2003). Anatomical mechanisms and functional implications of multisensory convergence in early cortical processing. *International Journal of Psychophysiology*, *50*, 5–17.
- Searby, A., & Jouventin, P. (2003). Mother-lamb acoustic recognition in sheep, A frequency coding. *Proceedings of the Royal Society of London B, Biological Sciences*, *270*, 1765–1771.
- Searby, A., Jouventin, P., & Aubin, T. (2004). Acoustic recognition in macaroni penguins, an original signature system. *Animal Behaviour*, *67*, 615–625.
- Shotter, J. (1996). Speaking bodies. *Theory and Psychology*, *6*, 177–179.
- Sidtis, D., & Kreiman, J. (2008). Let's face it, Phonagnosia happens, and voice recognition is finally familiar. In M. Pachalska & M. Weber (Eds.), *Neuropsychology and philosophy of mind in process. Essays in honor of Jason W. Brown* (pp. 298–334). Frankfurt/Lancaster: Ontos Verlag.
- Simonyan, K., & Jürgens, U. (2003). Subcortical projections of the laryngeal motor cortex in the rhesus monkey. *Brain Research*, *974*, 43–59.
- Stiles, W. B. (1999). Signs and voices in psychotherapy. *Psychotherapy Research*, *9*, 1–21.
- Stiles, W. B., Osatuke, K., Glick, M. J., & Mackay, H. C. (2004). Encounters between internal voices generate emotion, An elaboration of the assimilation model. In H. H. Hermans & G. Dimaggio (Eds.), *The dialogical self in psychotherapy* (pp. 91–107). New York: Brunner- Routledge.
- Terrazas, A., Serafin, N., Hernandez, H., Nowak, R., & Poindron, P. (2003). Early recognition of newborn goat kids by their mother, II. Auditory recognition and evidence of an individual acoustic signature in the neonate. *Developmental Psychobiology*, *43*, 311–320.
- Torriani, M. V. G., Vannoni, E., & McElligott, A. G. (2006). Mother-young recognition in an ungulate hider species, A unidirectional process. *The American Naturalist*, *168*, 412–420.
- Van Lancker, D. (1991). Personal relevance and the human right hemisphere. *Brain and Cognition*, *17*, 64–92.
- Van Lancker, D. (1997). Rags to riches, Our increasing appreciation of cognitive and communicative abilities of the human right cerebral hemisphere. *Brain and Language*, *57*, 1–11.
- Van Lancker, D., & Canter, G. J. (1982). Impairment of voice and face recognition in patients with hemispheric damage. *Brain and Cognition*, *1*, 185–195.
- Van Lancker, D., & Kreiman, J. (1986). Preservation of familiar speaker recognition but not unfamiliar speaker discrimination in aphasic patients. *Clinical Aphasiology*, *16*, 234–240.
- Van Lancker, D., & Kreiman, J. (1987). Unfamiliar voice discrimination and familiar voice recognition are independent and unordered abilities. *Neuropsychologia*, *25*, 829–834.

- Van Lancker, D., Cummings, J., Kreiman, J., & Dobkin, B. H. (1988). Phonagnosia, A dissociation between familiar and unfamiliar voices. *Cortex*, *24*, 195–209.
- von Kempelen, W. (1791). *Mechanismus der menschlichen Sprache nebst der Beschreibung seiner sprechenden Maschine (Mechanisms of human speech toward a description of a speaking machine)*. Vienna: J.B. Degen.
- von Kriegstein, K., & Giraud, A.-L. (2004). Distinct functional substrates along the right superior temporal sulcus for the processing of voices. *NeuroImage*, *22*, 948–955.
- Werner, H. (1956). Microgenesis and aphasia. *Journal of Abnormal and Social Psychology*, *52*, 347–353.
- Winer, J. A., & Lee, C. C. (2007). The distributed auditory cortex. *Hearing Research*, *229*, 3–13.

Diana Sidtis (formerly Van Lancker) is Professor of Communicative Sciences and Disorders at New York University and performs research at the Nathan Kline Institute for Psychiatric Research. An experienced clinician, her publications include numerous scholarly articles and book chapters.

Jody Kreiman is Professor of Head and Neck Surgery at the University of California, Los Angeles. She is a Fellow of the Acoustical Society of America and has published extensively in a variety of scientific journals and books.