

Review

Between sound and perception: reviewing the search for a neural code

Jos J. Eggermont *

Neuroscience Research Group, Departments of Physiology and Biophysics, and Psychology, University of Calgary, 2500 University Drive N.W., Calgary, AB, Canada T2N 1N4

Received 6 December 2000; accepted 29 January 2001

Abstract

This review investigates the roles of representation, transformation and coding as part of a hierarchical process between sound and perception. This is followed by a survey of how speech sounds and elements thereof are represented in the activity patterns along the auditory pathway. Then the evidence for a place representation of texture features of sound, comprising frequency, periodicity pitch, harmonicity in vowels, and direction and speed of frequency modulation, and for a temporal and synchrony representation of sound contours, comprising onsets, offsets, voice onset time, and low rate amplitude modulation, in auditory cortex is reviewed. Contours mark changes and transitions in sound and auditory cortex appears particularly sensitive to these dynamic aspects of sound. Texture determines which neurons, both cortical and subcortical, are activated by the sound whereas the contours modulate the activity of those neurons. Because contours are temporally represented in the majority of neurons activated by the texture aspects of sound, each of these neurons is part of an ensemble formed by the combination of contour and texture sensitivity. A multiplexed coding of complex sound is proposed whereby the contours set up widespread synchrony across those neurons in all auditory cortical areas that are activated by the texture of sound. © 2001 Elsevier Science B.V. All rights reserved.

Key words: Neural representation; Neural transformation; Neural coding; Auditory system; Neural synchrony; Amplitude and frequency modulation; Voice onset time; Speech; Vocalization

1. Introduction

Approximately 35 years after the publication of Kiang's monograph on the discharge patterns of auditory nerve (AN) fibers in the cat (Kiang et al., 1965) the physiology of the AN is well known. An important aspect thereof is the identification of the targets of the

myelinated type I and unmyelinated type II fibers in the three subdivisions of the cochlear nucleus (CN) (Ryugo, 1992; Liberman, 1991, 1993). For the type I fibers, the frequency tuning curves (FTCs), period histograms and post-stimulus time histograms (PSTH) for simple stimuli, e.g. clicks and tone bursts, are well documented. The responses to more complex stimuli, such as ele-

* Tel.: +1 (403) 220-5214 (office), +1 (403) 220-7747 (lab); Fax: +1 (403) 282-8249; E-mail: eggermon@ucalgary.ca

Abbreviations: AAF, anterior auditory field; AI, primary auditory cortex; AII, secondary auditory cortex; ALSR, average localized synchronized rate; AM, amplitude modulation; AN, auditory nerve; AVCN, anterior part of the VCN; BMF, best modulating frequency; CCI, consonant closure interval; CF, characteristic frequency; CN, cochlear nucleus; CNS, central nervous system; CV, consonant vowel; DCN, dorsal cochlear nucleus; DS, directional sensitivity; EP, evoked potential; ERP, event-related potential; FM, frequency modulation; FRA, frequency response area; FTC, frequency tuning curve; GBC, globular bushy cell; HVC, hyperstriatum ventrale pars caudale; IBE, information-bearing element; IBP, information-bearing parameter; IC, inferior colliculus; ICC, central nucleus of the inferior colliculus; ICX, external nucleus of the inferior colliculus; IHC, inner hair cell; ILD, interaural level difference; ITD, interaural time difference; LFP, local field potential; LIN, lateral inhibitory network; LL, lateral lemniscus; LSO, lateral superior olive; MGB, medial geniculate body; MGBv, ventral part of the MGB; MNTB, medial nucleus of the trapezoid body; MSO, medial superior olive; MTF, modulation transfer function; MU, multi-unit; MUA, multi-unit activity; P, posterior auditory field; PSTH, post-stimulus time histogram; PVCN, posterior part of the VCN; RCF, rates of change of frequency (velocity); RF, response field; rMTF, rate modulation transfer function; SBC, spherical bushy cell; SC, superior colliculus; SOC, superior olivary complex; SPL, sound pressure level; SR, spontaneous firing rate; STRF, spectro-temporal receptive field; tMTF, temporal modulation transfer function; VCN, ventral cochlear nucleus; VNLL, ventral nucleus of the lateral lemniscus; VOT, voice onset time; VP, ventral posterior auditory field; VS, vector strength

ments of speech, can generally be predicted from those to more simple ones such as tones, two-tone combinations and clicks (Sachs, 1984). Within 10 years after Young's (1998) detailed review on the CN, we perhaps will be at the same level of understanding about the CN, but the multiplicity of cell types and circuitry (Rhode, 1991; Oertel, 1999) causes this to be a more difficult endeavor. The ventral CN (VCN) extracts and enhances the frequency and timing information that is multiplexed in the firing patterns of the AN fibers, and distributes the results via two main pathways: the sound localization path and the sound identification path. The anterior part of the VCN (AVCN) mainly serves the sound localization aspects and its two types of bushy cells provide input to the superior olivary complex (SOC), where interaural time differences (ITDs) and level differences (ILDs) are mapped for each frequency separately (Carr, 1993). The posterior part of the VCN (PVCN) extracts across frequency timing aspects through its broadly tuned octopus cells, whereas its stellate cells, as well as those from the AVCN, compute estimates of the spectral representation of sound. This temporal and spectral information is carried, via the monaural nuclei of the lateral lemniscus (LL), to the central nucleus of the inferior colliculus (ICC). This sound identification path carries a level tolerant representation of complex spectra (e.g. of vowels) created by the chopper (stellate) neurons in VCN (May et al., 1998). The temporal and the spectral aspects of sound are both topographically, but mutually orthogonal, mapped in the ICC (Langner, 1992). The output from the SOC also arrives at the ICC, following some additional elaboration by the neurons in the dorsal nucleus of the LL (Wu and Kelly, 1995). In the ICC, ITDs and ILDs are combined into frequency-specific maps of interaural differences (Yin and Chan, 1988). Combining the frequency-specific ITD and ILD maps from the ICC results in a map of sound location in the external nucleus of the inferior colliculus (ICX). This auditory space map is subsequently represented in the deep layers of the superior colliculus (SC) (Middlebrooks, 1988) and aligned with the retinotopic map of visual space and the motor map of gaze (Hyde and Knudsen, 2000; Knudsen et al., 1987).

The inferior colliculi (ICs) and SCs form an important endpoint of the time-specialized part of the auditory nervous system (Trussell, 1997). In the IC, topographic maps are found for frequency, periodicity, and location of a sound. This spatial map is sufficient and necessary for adequate orientation to a sound source (Cohen and Knudsen, 1999). The ICC is the first level where physiological correlates of critical bandwidth properties, such as its level independence, are present (Ehret and Merzenich, 1988a,b; Schreiner and Langner, 1997).

In the last two decades one has slowly started to accept the notion that the auditory system evolved to allow the perception of sounds that are of survival value, and that the auditory system therefore has to be studied using such sounds. This neuroethological emphasis (Ohlemiller et al., 1994) has brought us a major understanding of the brains of auditory specialists such as echo locating bats (Suga, 1988, 1996) and the barn owl (Konishi et al., 1988). Speech is not fundamentally different, in the acoustic sense, from animal vocalizations, albeit that it is not as stereotyped (Suga, 1992). Human speech cannot have carried any weight in the evolutionary pressure that led to vertebrate hearing, and thus one cannot expect any particular selectivity and sensitivity in the vertebrate auditory systems, including humans, to speech. However, one may assume that human speech developed according to the constraints posed by the auditory and vocalization systems. This is reflected in the fact that human speech and animal sounds, not only those of other primates, share the same three basic elements: steady-state harmonically related frequencies, frequency modulations (FMs) and noise bursts (Mullennix and Pisoni, 1989; Suga, 1988, 1992; Fitch et al., 1997).

Most of our knowledge about the auditory thalamus and auditory cortex is obtained by stimulating with those sound elements that speech and animal vocalizations have in common: noise bursts, FM, and harmonic complexes interspersed with silent gaps. Frequency, noise burst bandwidth and preference for FM appear to be topographically mapped in cortical areas of auditory generalists (Schreiner, 1995) and specialists (Suga, 1994) alike. In contrast, sensitivity to low-frequency amplitude modulation (AM), as well as to gaps and voice onset times (VOTs), appears to be distributed across most neurons in at least three cortical areas and reflected as modulations in firing rate, that are synchronous across areas (Eggermont, 1994a, 1998c, 1999, 2000a). Many auditory cortical areas are tonotopically organized (Phillips, 1995; Schreiner, 1995) and they are presumably specialized to represent a limited, and likely different, set of particular important sound features (Ehret, 1997), albeit that none of these specializations, except in the mustache bat's cortex, has definitively been identified. One would expect that separate auditory cortical areas needed to be able to integrate biologically important sound features with other perceptual and cognitive tasks. It is therefore likely that individual cortical areas fulfill a role similar to that of the various cell types and subdivisions in the CN and brainstem. It seems likely that no more than a few independent channels or types of processing can coexist within an area (Kaas, 1987). The information, extracted by each cortical area, could be used to create clustered representations of sound location (e.g. in the frontal eye

fields, Cohen and Knudsen, 1999), and sound meaning (e.g. in the mammalian equivalents of premotor areas such as hyperstriatum ventrale pars caudale (HVC) in birds and Broca's area in humans, Doupe and Kuhl, 1999).

The neural code employed by a sensory system is likely determined by the innate structure of the system, i.e. it is the connection of the pathways and the properties of their neurons that produce the coded representation (Ehret, 1997). These anatomical connections and their neuronal specialization determine what kind of neurophysiological representation of the stimulus will occur. At higher levels of the central nervous system (CNS) these representations will be modulated by neural activity reflecting internal states such as drowsiness or arousal and also by the recent perceptual history of the animal (Merzenich and deCharms, 1996).

Perkel and Bullock (1969) asked more than three decades ago: 'Is the code of the brain about to be broken?' They subsequently listed a large number of potential neural codes that 'made sense' to the neuroscientists of that time. As we will see 'making sense' or 'having meaning' is crucial to the notion of a code; it indicates that coding occurs in context. The endless list of potential codes in that review also suggested that the concept of 'code' was very broadly defined. The specific mentioning of the 'code of the brain' suggests a belief that there is only one neural code for all perceptual phenomena.

Maybe we know more about the neurophysiological substrates of the vowel /*ɛ*/ than of any other speech sound. We know nearly everything about the neural activity evoked by the vowel /*ɛ*/ in AN fibers (Delgutte, 1984; Sachs, 1984) and in various cell types of the CN (May et al., 1998). On the basis of those neural responses the investigator is able to identify that vowel from a selection of other vowels with near certainty. The auditory CNS can do so too, but lacks the *a priori* knowledge of the experimenter about which vowels are presented. How does the CNS do this identification? The neural responses to /*ɛ*/ will likely change dramatically between CN and auditory cortex. Is there, in the end, a unique code for /*ɛ*/? If so, how is that code formed out of the neural activities evoked by the sound? The vowel /*ɛ*/ can be characterized by a unique spectrum that appears to be represented in a population of T-multipolar cells (choppers) in VCN in a level tolerant way (Young, 1998). It most likely can also be uniquely represented in the population autocorrelogram of the phase-locked firings in a population of bushy cells or octopus cells in VCN (Oertel, 1999). However, we do not have any account of neural activity caused by /*ɛ*/ in, for instance, the IC or the auditory cortex. We

do know the multi-unit activity (MUA) that is produced in some parts of auditory cortex by other phonemes such as /*da*/, /*ta*/, /*ba*/, and /*pa*/ (Steinschneider et al., 1994; Eggermont, 1995a,b). We even know what areas in the human brain are metabolically activated differentially by presentation of voices and other sounds through visualization by positron emission tomography scans or functional magnetic resonance imaging (Belin et al., 2000). However, vowel representation will strongly depend on context, i.e. what consonants precede or follow it. It is therefore not clear if a representation of a word can be generated from the representation of its phonemes in isolation.

There is evidence that different acoustic representations exist for identical phonemes. Identical cues for particular phonemes can also give rise to different percepts as a function of context. As a result of these contextual effects, it has been difficult to isolate acoustic attributes or features that satisfy the perceptual invariance that is observed in practice. Thus, there is no simple one-to-one relationship between the acoustic segments of the speech waveform and the way they are perceived (Mullennix and Pisoni, 1989). The implications thereof for the existence of a neural code have not yet been explored.

The relative importance of spectral and temporal information present in the sounds used to convey speech understanding has been elucidated by research in cochlear implant users where limited place-specific information is delivered to the remaining AN fibers. What minimal information has to be presented to the AN fibers in case of complete sensory hearing loss so that the receiver of a cochlear implant can fully understand speech? It seems that, at least under optimal listening conditions, spectral information is far less important than temporal information for recognition of phonemes and words in simple sentences (Shannon et al., 1995). Event-related potential (ERP) research in normal hearing and implanted human subjects has elucidated a potential role for non-primary auditory pathways in signaling temporal differences in sounds that can or cannot be sensed by the auditory cortex (Ponton et al., 2000).

This review will present an extensive, albeit not exhaustive, selection of what is known about neural responses in the auditory system of laboratory animals, such as monkeys, cats and bats, related to the identification of complex sound. I will discuss the transformations in the neural activity that take place along the way from AN to cortex. The review will also speculate further, extending a previous review (Eggermont, 1998b), on that elusive interface between stimulus and perception: the neural code.

2. Biologically important features of sound have shaped the auditory nervous system

Worden and Galambos (1972) aptly noted that ‘the full capacity of the sensory processor may not be revealed except through study of its response to stimuli that pose analytical tasks of the kind that shaped its evolutionary development’. Thus, regularities in the acoustic biotope, consisting of individual vocalizations and background sounds that are part of the natural habitat of the animal (Aertsen et al., 1979; Nelken et al., 1999; Smolders et al., 1979), are likely manifested in the response properties of auditory neurons. It has been suggested that the statistical structure of natural signals is important for creating efficient representations of the sensory world in the CNS (Barlow, 1961; Ruderman and Bialek, 1994; Rieke et al., 1995). One can also say that sensory systems have evolved to efficiently process natural stimuli. In this respect, the peripheral auditory system may have evolved not only for the reception of biologically important sounds, but also for the frequency–time analysis of those sounds that are especially important to a species. In social animals, the processing of communication calls is likely one of the major functions of the central auditory system. Spectrograms of speech and animal vocalizations (Fig. 1) all exhibit three basic components (information-bearing elements, IBEs; Suga, 1989): constant frequency parts, noise bursts, and FM components. In human speech, several harmonically related constant frequencies (formants) constitute a vowel, the transition from a noise burst into the vowel is typically in the form of a formant glide, an FM part (Mullenix and Pisoni, 1989; Fitch et al., 1997). Combinations of the IBEs are important for affecting the behavior of both animals and humans. The combinations of IBEs are frequently stereotyped in animal vocalizations but vary enormously in human speech. In fact, one of the most fundamental characteristics of speech is this inherent physical variability while keeping its perceptual constancy.

Auditory information is carried not only by the acoustic parameters characterizing each of the above three types of IBEs, but also by information-bearing parameters (IBPs) representing relationships among these three IBEs in the frequency, amplitude and time domains (Suga, 1992, 1996). Examples of IBPs could be the characteristic delay between the noise burst and the FM component, the VOT, in some phonemes. Another example is the rate of the simultaneous AM of constant frequency components in the vowel that, depending on modulation frequency, represents glottal pulse rate and prosody.

One could tentatively equate the IBEs with the texture of sound, and the IBPs with the contours of sound.

Texture is largely spectral, whereas contours are largely if not exclusively temporal in nature.

Speech and animal vocalizations contain frequency-modulated components in which frequency and time are not separable, that is, the product of the spectrum and the temporal envelope is not equal to the spectrogram of the original vocalization (Eggermont et al., 1981). In contrast, mixtures of vocalizations tend to be separable into a stationary Gaussian process (the carrier) and a temporal envelope. Thus all frequency components in the carrier are co-modulated by the envelope. This low-frequency co-modulation of background sound may be acquired during atmospheric propagation of sound as a result of micro turbulence (Richards and Wiley, 1980). Co-modulation of the background makes it less effective in masking individual, foreground, vocalizations which are not subject to this co-modulation (Nelken et al., 1999).

Small groups of neurons may have evolved to process sounds that are essential for species survival and may bind, i.e. simultaneously represent, many IBPs of these sounds. On the other hand, neurons or neural populations that show little binding may process sounds that are not particularly important to the species. In other words, the processing of these general sounds may be solely based upon their spatio-temporal patterns (Suga, 1994). This could imply that a neural code exists only for those sounds that are essential to a species but not for sound in general. However, what is essential for survival may change during the life span of an animal and could for instance move sounds from the, initially, general category into the, presently, essential category. So context dependence or learning with respect to the information carried by biological sounds cannot be neglected, especially when taking brain plasticity into account (Merzenich and deCharms, 1996). The complexity and variability of vocal signals and the processes devoted to perceiving the signal constitute an important difference between humans and animals in perceiving complex sounds. However, the auditory aspects of speech processing may depend on mechanisms specialized for the representation of temporally complex acoustic signals in humans and other mammals alike, regardless of their communicative or linguistic relevance (Fitch et al., 1997).

3. Representations and codes: defining the neurophysiological emphasis

I propose that in the field of neurophysiology, the concept ‘neural code’ be reserved for a unique specification, a set of rules, that relates behavior to neural activity. The neural code is then to be equated with the result of the last non-identity transformation of

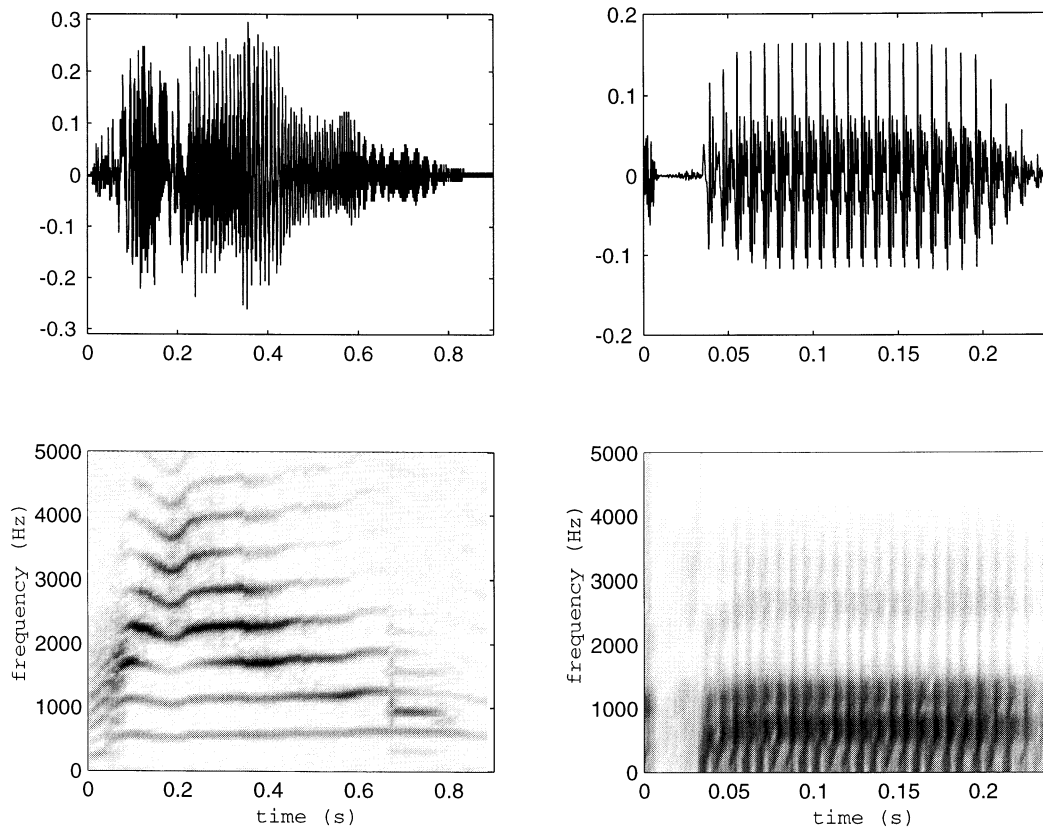


Fig. 1. Two vocalization sounds that illustrate similarities and differences in IBEs and IBPs. In the left-hand column, the waveform and spectrogram of a kitten meow are presented. The duration of this meow is 0.87 s, the average fundamental frequency (F_0) is 550 Hz, the lowest frequency component (F_1) is about 0.5 kHz and the highest frequency component (not shown) is 5.2 kHz. The second (F_2) and the third (F_3) harmonics, between 1.5 and 2.5 kHz, have the highest intensity. Distinct downward and upward FMs occur simultaneously in all formants between 100 and 200 ms after onset. The meow has a slow AM. In the right-hand column, the waveforms of a /pa/ syllable with a 30 ms VOT and its spectrogram are shown. The periodicity of the vowel and the VOT are evident from the waveform. Low level aspiration noise was present in the period before the onset of voicing. The dominant frequency ranges are F_0 = 125 Hz, F_1 = 700 Hz, F_2 = 1200 Hz. Because the dynamic range of the representation is only 30 dB the third formant at 2600 Hz is only weakly visible. The fundamental frequency started at 125 Hz and remained at that value for 100 ms and dropped from there to 100 Hz at the end of the vowel. The first formant started at 512 Hz and increased in 25 ms to 700 Hz, the second formant started at 1019 Hz and increased in 25 ms to 1200 Hz and the third formant changed in the same time span from 2153 Hz to 2600 Hz.

neural activity between sensory receptor and cortex. In this sense it may be analogous to the genetic code present in DNA (words of three nucleic acid bases, out of four available, code for specific amino acids) that is translated by RNA (through the ribosomes) into proteins (combinations of 20 amino acids) which are the phenotype (Frank-Kamenetskii, 1997). From a cryptographic standpoint, however, the genetic code is not a code. The string 'genetic code' is merely a metaphor for the correlation between nucleic and amino acids. Similarly, the 'neural code' could be defined as not more than the neural basis of a correlation between stimulus and behavior.

If a neural code such as defined above exists, how does it materialize into its phenotype: behavior? In this I include perception among behavior. One could wonder if there is also a cognitive tagging process that helps recognizing parts of the code in different contexts? Just

as the genetic environment is important to determine which genes (combinations of nucleic acid 'words') are expressed and at what time, there is a context dependence that determines if and when the neurally coded stimulus elicits a behavior. Could one also say that the same behavior (percept) might be elicited by different neural codes just as different DNA 'words' may code for the same amino acid?

The labels 'neural code', coding, decoding and encoding all have been used very widely and most of the times were used synonymously with neural representation. Some object entirely to the use of these labels on the basis that 'decoding of sensory input, followed by representation and reconstruction of the world within the brain, logically requires that these rebuilt images, assembled from the previously encoded sensory input, must now be viewed or processed by neural structures that do not require encoded input' (Halpern, 2000).

However, this may be too strict an interpretation of these coding-related labels. The idea behind encoder and decoder operations in communication systems is that the signal is transmitter as a sequence of symbols. The relationship between the signal and the symbols, generally not a one-to-one relationship, is usually called the code. Part of the problem in applying this to perception is that for the nervous system, the set of symbols itself is unknown. The symbols could be spike times, spike time intervals, number of spikes, particular spike sequences, spike time coincidences, etc.

Johnson (1980) and Cariani (1995) have extensively addressed the necessary distinction between representation and code in a general context, and it is useful to review and potentially extend their classification. It is also useful to state up front that what the auditory system up to at least the level of the primary auditory cortex (AI) does is process information about sounds. Information that is not present at the input of a specific nucleus cannot be present at its output, and neural coding may be a unique way to reflect the essential, abstracted, information leading to behavior. This view bypasses cognitive substitutions, i.e. making assumptions about and identification of a stimulus that are not warranted on the basis of its physical attributes.

A discussion of information processing requires that distinctions are made between its three interdependent but separate aspects: neural representation, neural transformation, and neural coding.

The neural representation of a sound can be defined as the complete spatio-temporal array of neural activity in all of the neurons passing through a transverse plane across the auditory system (Johnson, 1980). This plane can be at the level of the AN, the brainstem, the mid-brain, etc. Cariani (1995) calls the neural correlate of stimulus quality a 'sign'. A sign is a characteristic pattern of neural activity that has all the information needed to affect a particular sensory discrimination or behavior. When the sign is present, the particular sensory discrimination or behavior may occur but it does not occur when the sign is not present. This is in many ways equivalent to the neural representation of a behaviorally meaningful sound. Thus, a sign is more restricted than a neural representation because, for instance, the neural representation for the vowel /e/ does not have much meaning for the experimental animal in which the recordings are done and thus is generally not a sign. However, after some behavioral training such a neural representation may become a sign.

The first 'neural' representation resides in the cochlear hair cells. Subsequently, this representation is conveyed by parallel pathways, embodied in the collaterals of the AN, to the three divisions of the CN, diverging further as outputs from distinct cell types and continuing toward the IC. This creates a set of parallel neural

pathways. A particular pathway may contain a detailed representation of the stimulus, but does not contribute to recognition of the stimulus. For instance, the localization pathway originating in the AVCN may not be used in stimulus recognition at all (but see Loeb et al., 1983; Koch and Grothe, 2000; Grothe, 2000). Alternatively, a particular pathway may carry only a partial representation of the sound, but may provide all of the information about particular important sound dimensions (IBEs and IBPs) that need to be discriminated. This all has to be taken into account when evaluating neural coding.

A neural transformation is the computational process by which a neural representation at one location in a sensory pathway is derived from a neural representation at a lower level. From a system's analysis point of view, the transformation between two neural representations can be expressed as a transfer function or as the action of a filter. For a linear transformation with independent inputs, the transformation could be described completely by a multi-dimensional impulse response. Transformations occur at synaptic junctions, and they are in general non-linear. Unless synapse specialization occurs (as it does in the AN–spherical bushy cell (SBC) or in the globular bushy cell (GBC)–medial nucleus of the trapezoid body (MNTB) principal cell synapses) those transformations could destroy synchrony to the stimulus fine structure and may, for example, set the limits for the discrimination of rapidly changing stimuli. One generally loses information by neural transformations because most transformations create noise. The only way to gain is when the conditional probabilities for the response given the sounds do change, or when the noise is reduced, e.g. as the result of converging neural activity such as occurring in the octopus or chopper cells of the CN. Note that the gain in these cases is in an individual output neuron with respect to the information carried by the activity of individual neurons in the input representation. Nothing can be gained in the representation as a whole.

Neural code has been difficult to define. Some proposals: '...transformed and transmitted signals do not constitute a bona fide neural code unless that information is appropriately acted upon – interpreted – by appropriate parts of the organism' (Perkel and Bullock, 1969). '...we shall operationally define a true code as a parameter of the signal which actually carries behaviorally usable information' (Uttal, 1969). '...a neural code at a particular location within the sensory pathway as being the parameter of the afferent discharge in the population that is actually used by the organism in particular sensory behavior' (Mountcastle, 1975). 'Codes are the functional organizations that actually utilize a particular set of signs to effect a perceptual discrimination' (Cariani, 1995). All cited sources agree

that a code relates neural activity to behavior. Brugge (1992) states that ‘in order to qualify as a neural code for acoustic information, it must be shown first that the (neural) pattern in question occurs in the auditory system under natural conditions or is evoked by natural stimuli, and second that there exists a sensitive receiver; that is a set of neurons whose activity changes in response to the candidate code it receives’. This definition is close to that of a ‘sign’, a specific form of a neural representation.

Again adopting a system-theoretic point of view, the neural code could be defined as the transfer function between a (multi-dimensional) stimulus and behavior (perception). The maximum information that can be transmitted by the neural code from this point of view is related to the signal to noise ratio, i.e. the ratio of stimulus-induced spikes and spontaneous activity. The latter should ideally be independent and additive to the stimulus-induced activity, however, this does not appear to be the case (Cecchi et al., 2000). A problem is that the stimulus needs to be defined as all relevant sensory information that results in the behavior or percept. A further complication for a systems approach is that the nervous system is not time-invariant (is changed by experience), is not deterministic (no unique relationship between stimulus and percept), and certainly is not linear.

Representations in cortex may be relational, i.e. relations between elements and ensembles could be used to establish reliable and flexible representations. This is a way in which the representational relations among a group of neuronal elements, modified by learning and plasticity, can remain invariant (Merzenich and deCharms, 1996). Relational representations could be at the basis of the maintenance of perceptual constancy in a changing cortex. Thus, neural representations are not static: one may learn to attend to some components of the neural representation and ignore others (the ‘selective listening’ hypothesis). This could be represented by an adjustment of the neural code; the rule connecting stimulus and response. Alternatively, one could assume that the neural code is unique and that the read out of the neural code is changing dependent on the context. I prefer this latter approach because it keeps the analogy with the unique genetic code where the read out of the gene (expression) depends on its environment. On the down side, a neural code defined this way could be intractable by neurophysiological techniques. The neural code provides a link between a neural representation, based on neurophysiological studies, and sensory discrimination performance (a behavior) that depends on the information contained in that neural representation (deCharms and Zador, 2000). Thus I intend to isolate the content of the neural code from its ultimate function as a releaser of behavior, thereby presuming

that it can release a specific behavior or cause a specific percept. Whether the code actually elicits the behavior will not be considered here. This restriction, in principle, allows a limited search for the neural code in anaesthetized or otherwise non-behaving animals. Later on I will come back to the system-theoretic approach and interpret the neural code in terms of the mutual information (Borst and Theunissen, 1999) present in a cortical neural representation.

There is a hierarchical order in this classification of neural activity: a neural representation forms the basis, a sign is a neural representation of behaviorally meaningful sounds, and a code is a sign that on its own is capable of evoking a behavior. This definition limits the use of neural code to the final interface between sensory neural activity and behavior; an interface that likely resides in cortex. Thus, it is probably meaningless to speak of complex sound coding in the discharges of the AN fibers. On the other hand, at the auditory cortex level neural representations and neural codes may both exist. It is not a priori clear at what level in the auditory nervous system one can justifiably talk about neural coding, but it will be argued that it cannot be at a level below that of the IC.

How does the neural code relate to the perceptual distinctions that can be made? Distinctions in perception are based upon distinct activity patterns of the nervous system, either specifically sensory or those reflecting the animals internal state. Thus each dimension of perceptual quality should be related to a dimension of neural activity. It should thus be possible to correlate spaces of perceptual distinctions with spaces of neurally encoded distinctions (Cariani, 1995). The correlation matrix that contains this information, reflects, or is identical with, the neural code.

4. Stimulus representation in terms of firing rate or firing times

Usually one reads about rate coding, or about temporal or synchrony coding, in the activity patterns of the AN. In our restricted nomenclature this is translated into the representation of a specific sound in the overall or synchronized firing rates of AN fibers or in the interspike intervals of AN fiber activity. In the following I will transcribe part of Cariani’s (1995) classification for neural code types into those for neural representations.

Neural representations reflecting the connectivity pattern of neurons and nuclei take the form of place, labeled line, or spatial pattern representations (Brugge, 1992). In the auditory system this usually takes the form of firing rate vs. characteristic frequency (CF) or phase-locked (synchronized) rate vs. CF (Young and

Sachs, 1979). A problem with a rate representation is that it rapidly deteriorates in the presence of background noise (Miller and Sachs, 1984). Neural representations may also reflect the temporal patterning of firing activity without the need for any specific reference to CF. Such a representation can be characterized by the (all-order) interspike interval representation of stimulus periodicity as demonstrated in AN fibers (Cariani and Delgutte, 1996), in the IC (Delgutte et al., 1998), and in auditory cortex (Eggermont, 1998c). In this form one can speak of a temporal representation of sound (see Theunissen and Miller (1995) for a more differentiated point of view). Neural representations can also reflect the relative time of arrival as quantified by latency differences as a function of a particular stimulus dimension (e.g. sound source azimuth, Eggermont, 1998a) or degree of interneuronal synchrony (deCharms and Merzenich, 1996; Eggermont, 1994b, 1997a, 2000a).

A main consideration here is to acknowledge that the stimulus and its time of presentation are unknown to the animal. Thus one of the problems in identifying a useful code or representation for the CNS is that such a code or representation cannot in any way incorporate properties of the external stimulus such as time of presentation, frequency content or position in space. For instance, one cannot use period histograms which require knowledge of the stimulus periodicity to infer timing representation in a periodic sound but one can use the neuron's own interspike interval distribution which presents the same information and needs no reset of an internal 'clock' (Horst et al., 1986; Javel et al., 1988). One cannot use latency as part of the stimulus-response features used in a 'panoramic code' (Middlebrooks et al., 1994) but one may use spike latencies relative to an internally available global time marker such as a local field potential (LFP). In practice, however, this does not seem to make much difference with respect to the representation of sound source location (Furukawa et al., 2000). However, internal representations or codes, in principle, need to be free from external 'anchors'.

These types of neural representation are by no means all inclusive. For instance, in the AN a firing rate representation of speech exists along side, and is multiplexed with, a synchronized firing rate representation (Sachs, 1984), and also with an interspike interval representation (Cariani and Delgutte, 1996; Cariani, 1999a). The cells in the CN that receive information from AN fibers may extract either the firing rate information (e.g. the stellate cells), or the firing synchrony information (e.g. the bushy cells and octopus cells). This form of multiplexing information in neural representations may be a consequence of a specialization for timing that exists in the subcortical auditory nervous system.

5. Neural representations as maps

A topographic map is defined as an ensemble of neurons combining two or more neural representations with parameters systematically ordered along a spatial dimension in a given nucleus or area. Usually this takes the form of a spatially coded parameter (e.g. CF) and some other parameter (e.g. average firing rate, first-spike latency) (Schreiner, 1995; Ehret, 1997). As an example, the rate-place scheme's central representation is formed by differences in spatially organized firing rates, i.e. as firing rates differing as a function of CF. At behaviorally relevant intensities, the rate tuning curves are usually broad, so that a relative firing rate representation (Erickson, 1974) has to be extracted from the topographic maps to match the often very fine perceptual capacity. An example is the map of sound source azimuth in the SC. Latency-place representations use the relative time of arrival in different spatial locations in order to encode intensity and other qualities. Latency-place mechanisms appear to be involved in a wide variety of sensory processes: e.g. stimulus localization, echo delay maps (Cariani, 1995, 1997, 1999b). In the AI of the bat, neurons are tuned for amplitude as well as for frequency of sound. The best amplitude map varies along a dimension perpendicular to the tonotopic axis. A higher order area (FM-FM area) in the bat's secondary auditory cortex (AII) maps the time interval between acoustic events such as emitted sound and received echo that signals prey distance. The best time interval between two FM sounds varies systematically across the cortical surface. Tonotopy is absent in the FM-FM area and neurons are grouped in clusters according to the ranges of frequencies contained in the FM components (Suga, 1994).

Besides topographic maps, where the similarity of the representational properties of two neurons is reflected in their physical distance, there are also functional maps. Such maps have an ordering through their correlation structure only (Koenderink, 1984). Neural units with strong correlation of their firings can be considered close in the neural organization, units with weak correlation have a larger functional distance. In visual cortex, distant neurons that have similar orientation preference tend to show significant correlation in their firing times (Ts'o et al., 1986), suggesting that these like-minded neurons form a correlation map of orientation preference that is superimposed in a fractionated way upon the retinotopic map. Maps of this kind have yet to be identified in the auditory cortex, where correlation strength appears relatively isotropically distributed (Brosch and Schreiner, 1999; Eggermont, 1992a, 1994b, 2000a).

The apparent importance of self-organizing processes

in development, based on correlated neural activity patterns, suggests that some specific features of cortical organization, such as tonotopic representations, and types of modular grouping of neurons within those representations, could be byproducts of developmental sequences in the building of brains. Thus, specific features of cortical organization may be the outcome of the building process rather than a design for optimizing function (Kaas, 1987).

Functional maps may differ for different stimulus conditions or different contexts, whereas topographic maps are generally situation-independent. However, one has to keep in mind that what is mapped topographically is often an experimenter-based abstraction, e.g. CF for the tonotopic map. Because CFs for different neurons are found at different threshold intensity levels, such a map may not have functional importance besides reflecting the developmental/experiential process that led to it.

A prime example of a functional map is that of auditory space in the ICX. Such maps are computed in hierarchical fashion from frequency-specific topographic maps of interaural time and intensity differences in the SOC. Inputs to the neurons that make up such a map are almost instantaneously transformed into a place-coded distribution of the mapped parameter. Sorted in this manner, the derived information can be accessed readily by higher order processors (mapped or not) using relatively simple schemes of connectivity (Knudsen et al., 1987).

In most computational maps, and also in some topographic maps, regardless of the parameter being analyzed, neurons are typically broadly tuned (Seung and Sompolinsky, 1993) for the mapped parameter. Computational maps are pre-wired and are fast but still modifiable. They are likely only created for parameter values that are biologically relevant and require speedy action. Despite the fact that neurons organized in computational maps are broadly tuned for the mapped parameter, precise information about the values of parameters is contained in the output of these maps. The neuron's tuning curves are peaked and shift systematically across the map. This gives rise to systematic differences in firing rates across the map for any given stimulus. Thus, high-resolution information is contained in the relative responses of neurons across the map. This constitutes a ratio map (Erickson, 1974) that is level tolerant if the rate intensity functions of the neurons are approximately linear. The subsequent processor has to be sensitive to the relative levels of activity within a large population of neurons, and able to detect locations of peak activity within the map. Computational maps, with their parallel array of preset processors, are ideally suited to rapidly sort and process components of complex stimuli and represent the results in

a simple systematic form. Regardless of the stimulus parameter processed, the answer is always represented as the location of a peak of activity within a population of neurons. When a parameter is represented in topographic form, a variety of neuronal mechanisms can operate to further sharpen tuning in ways not possible if the information is in a non-topographic code. One class of mechanisms is regional interactions such as local facilitation and lateral inhibition, which can only work on mapped information (Knudsen et al., 1987).

Maps provide for binding of IBEs and IBPs. This could be done if IBEs are topographically mapped and IBPs are providing a temporal tagging of the IBEs across maps. A sound ultimately forms a correlation map that can be considered as a neural assembly. A series of sounds may result into a series of concatenated maps. This is an ever-changing assembly that forms a path through the *N*-dimensional internal representation of the acoustic biotope.

There are multiple, often nearly identical, repetitions of the tonotopic map at both subcortical and cortical levels. Each separate frequency representation is a unit in a system of serial/parallel channels making up the central auditory system. In theory, the number of parameters that can be mapped independently and continuously in one area is limited to the number of dimensions in the neural structure (Kaas, 1987; Schreiner, 1995). The nervous system potentially overcomes this limitation by organizing fine-grained maps within coarse-grained maps, such as the visual orientation map within the retinotopic map. However, map-dependent neural interactions operate optimally only for the parameter mapped in fine grain. Perhaps this is one reason that different parameters are mapped in separate, functionally specialized areas of the brain, most of which also contain a coarse tonotopic map. Each map is likely performing a different type of analysis on the sensory information from the receptors; perception ultimately involves integration of the information from these separate representations (Young, 1997). This could be done on the basis of the underlying tonotopicity, but also on the basis of a synchronized representation of contours. Understanding the functional roles of the separate maps depends on anatomical evidence, physiological properties, and behavioral evidence after activating or deactivating the putative map.

Do topographic maps participate directly in neural coding? Specifically do topographic maps convey useful information? It is clear that tonotopic maps are empirical constructs defined by the investigator in terms of an arbitrary level of neuronal firing in response to stimulation of specific cochlear locations. In fact, the topography of cortical maps can change considerably with stimulus level (Phillips et al., 1994), localized sensory deprivation (Harrison et al., 1991; Rajan et al., 1993;

Eggermont and Komiya, 2000) and experience (Dinse et al., 1993). The basic organization of a sensory map is established through developmental processes that allow easy lateral interactions between neurons. Short-term changes in the map reflect modifications in the balance of excitatory and inhibitory inputs reaching certain neurons. However, maps may not provide direct information about profiles or levels of neuronal activity produced by different stimuli. Maps can be interpreted or ‘read out’ only by an external observer; they cannot be utilized internally by the brain, which ‘knows’ only neuronal activity or the lack thereof. This leads to the conclusion that information about different stimuli and their functional significance is not conveyed directly by the map structure but only indirectly by patterns of activity resulting therefrom.

6. Information processing: probing the efficiency in neural representations

In general, calculating whether the information needed to represent a particular stimulus property is present in the firings of a given neural population is much easier than to determine whether the CNS actually utilizes all or part of this information to modify its behavior (Johnson, 2000). A clear example is found in the relationship between AN activity and the threshold of hearing. In barn owls, the threshold for phase locking, i.e. being able to tell the frequency of the sound from the representation in the interspike intervals, is between 10 and 20 dB lower than the threshold at which the firing rate begins to rise. The threshold of firing rate increase, in fact, corresponds to the threshold of hearing at that frequency (Köppl, 1997). In somatosensory cortex, thresholds based on firing rate were also similar to the animal’s psychophysical threshold whereas neural thresholds based on periodicity were far lower than those behavioral thresholds (Hernandez et al., 2000). Earlier studies in cat and chinchilla (Javel et al., 1988) also found a lower threshold for changes in phase locking compared to firing rate. However they suggested that there was a better correspondence between behavioral hearing thresholds and those based on the emergence of significant phase locking. So there is clearly information about the sound in the temporal patterning of neural activity, sometimes well below the threshold of hearing is reached. Either that information is lost along the auditory pathway, or the decision about hearing or not hearing is only based on a noticeable increase in firing rate. Neurons in auditory cortex, however, can signal the presence or absence of a sound better on the basis of firings synchronized across neurons in different cortical areas than on increases in firing rate (Eggermont, 2000a).

Information theory, the most rigorous way to quantify the content of neural representations, is an aspect of probability theory that was developed in the late 1940s by Shannon (1948) as a mathematical framework for quantifying information transmission in communication systems. Shannon’s information theory can be considered as a generalized measure of stimulus discriminability. Given a, a priori known, set of stimuli, the amount by which the uncertainty about the stimulus $H(S)$ is reduced after observing a response $H(S/R)$ is called the mutual information:

$$I(R, S) = H(S) - H(S|R) \quad (1)$$

In general, the mutual information can be written as a function of both the conditional and marginal stimulus and response probabilities:

$$I(R, S) = -\sum_i p(S_i) \log(p(S_i)) - \sum_j p(R_j) \sum_i p(S_i|R_j) \log(p(S_i|R_j)) \quad (2)$$

and can be expressed in bits/spike. Here $p(S_i)$ is the a priori distribution of stimulus parameter S_i (the one to be discriminated), and $p(S_i|R_j)$ is the a posteriori conditional distribution of the stimulus parameter S_i for a given response R_j . This equation implies that in order to evaluate the conditional entropy $H(S/R)$, one first has to convert back from the neuronal response distribution, to a stimulus probability distribution by means of Bayes’ rule:

$$p(S|R)p(R) = p(R|S)p(S) \quad (3)$$

This is easy for constant stimulus sets (e.g. containing nine morphed cat vocalizations, Gehr et al., 2000) but not for natural stimulus ensembles. The stochastic mapping of the sensory environment S onto the set of neural activity patterns R , $p(R/S)$, forms the neural representation of the stimuli that make up the sensory environment. This is what is usually studied experimentally. The inverse mapping $p(S/R)$ gives the plausibility of a sensory stimulus S given the recorded neural activity R , and can be considered as the sensory interpretation of neural activity (Johannesma, 1981). This is what the animal uses to navigate in perceptual space. We have called these two approaches previously experimenter-centered and subject-centered approaches (Eggermont et al., 1983a,b). The analysis of neurophysiological data from the point of view of the subject is generally done via the reverse correlation method (de Boer, 1967, 1969; de Boer and de Jongh, 1978). In a specific implementation this may result in the spectrotemporal receptive field (STRF), approximately equal

to the average spectrogram of sounds preceding a spike (Eggermont et al., 1983a,b; Kim and Young, 1994; deCharms et al., 1998; Theunissen and Doupe, 1998). This reverse correlation approach glorifies the information residing in the ‘single spike’, rather than in the population activity (see also Rieke et al., 1997). The result is either the average preferred waveform, the reverse correlation function (in case there is phase locking of discharges to the sound), or the average preferred frequency–time distribution of sounds preceding the spikes. Such preferred sounds can be considered optimal in terms of being matched to the neurons spatio-temporal response properties (deCharms et al., 1998), but they do not necessarily evoke the highest firing rates.

A universal finding in information calculations carried out for peripheral neurons is a relatively high importance of single spikes (Rieke et al., 1997) in the sense that the mutual information per spike is high. Information measures in bits per spike do not translate directly into spike timing precision in ms, but in the linear case, the high-frequency cut-off of stimulus encoding corresponds to the limiting accuracy of spike timing. In general, the auditory periphery and brainstem are highly specialized to process timing information.

Several pertinent questions have to be answered. Can $p(S)$ be defined in general? Does this definition require inclusion of the entire acoustic biotope with all frequencies of occurrence for all the individual sounds or sound combinations? Does this have to be weighted by the meaning of the sound, e.g. its meaning for reproduction or survival? Can mutual information change by learning? Is this reflected in a change in $H(R/S)$? Could one redefine the probability of occurrence of a stimulus $p(S)$, and thus $H(S)$, by its probability of impact but leave $H(R/S)$ untouched? This is not practical in most experiments.

Similarly $H(R)$, the uncertainty about (or entropy of) the neural response, corresponds to the number of bits required to specify all possible responses under all possible stimulus conditions. $H(S/R)$ is the entropy remaining in the stimulus once the neural responses are known. Adding the uncertainty remaining in the neural response when the stimulus is known, the neuronal noise $H(R/S)$, to $I(R,S)$ gives the total neural entropy, $H(R)$. Therefore an alternative expression for the mutual information is:

$$I(R, S) = H(R) - H(R|S). \quad (4)$$

Because $H(R)$ represents the maximal information that could be carried by the neuron being studied, comparing $H(R/S)$ to $H(R)$ also gives an estimate of the neural code’s efficiency.

The amount of information present in the spike train

can also be estimated by a comparison of the reconstructed stimulus and the original stimulus. Stimulus reconstruction can be done by, e.g., replacing each spike in a spike train by its reverse correlation function (Johannesma, 1981; Gielen et al., 1988) or by substituting the STRF (Hesselmans and Johannesma, 1989; Theunissen et al., 2000). The lower bound of the information present in the spike train can then be obtained from integrating $-\log(1-\gamma^2)$ across frequency, whereby γ^2 is the squared coherence between the reconstructed signal and the original (Borst and Theunissen, 1999).

Estimating information transfer without making assumptions about how the stimulus is encoded is done in the direct method. The ‘direct method’ calculates information directly from the neural response by estimating $H(R)$ and $H(R/S)$, it estimates exactly the average information transmitted but does not reveal what aspects of the stimulus are being encoded. This suggests that one can evaluate information rates by using two types of stimuli. (1) For the evaluation of the first term $H(R)$, total spike train uncertainty, one needs a large range of stimuli drawn randomly from a given stimulus ensemble, because these stimuli have to test the limits of response variability. (2) The second term $H(R/S)$ reflects the variability of the response when the stimulus is fixed, and can be evaluated from responses to a typical stimulus instance repeated many times. These newly developed information-theoretic methods (Buracas and Albright, 1999) allow one to quantify the degree to which neuronal representations become more abstract, by selective loss of irrelevant information, as one proceeds from the periphery to the CNS. It appears that most information carried by spikes is about the timing of abrupt variations in the stimulus, i.e. about the stimulus contours. It may thus well be that the neural code will be mostly reflecting a sequence of changes in neural activity.

7. Parallel distributed processing and a specialization for representing time characterize the auditory system

Sound is special among sensory stimuli. Sound source localization, in contrast to the visual and somatosensory system where stimulus location is directly mapped onto the receptor surface, has to be computed from interaural spectral and timing differences. The position of a sound source produces only minute time of arrival differences at the two ears: in the human ear at most 800 μ s, for a sound located along the axis through the ears. On that basis, we are able to distinguish differences in location that produce about 10 μ s interaural time difference (corresponding to a path length difference of approximately 3.5 mm). The specialization of the auditory system for the accurate processing of time

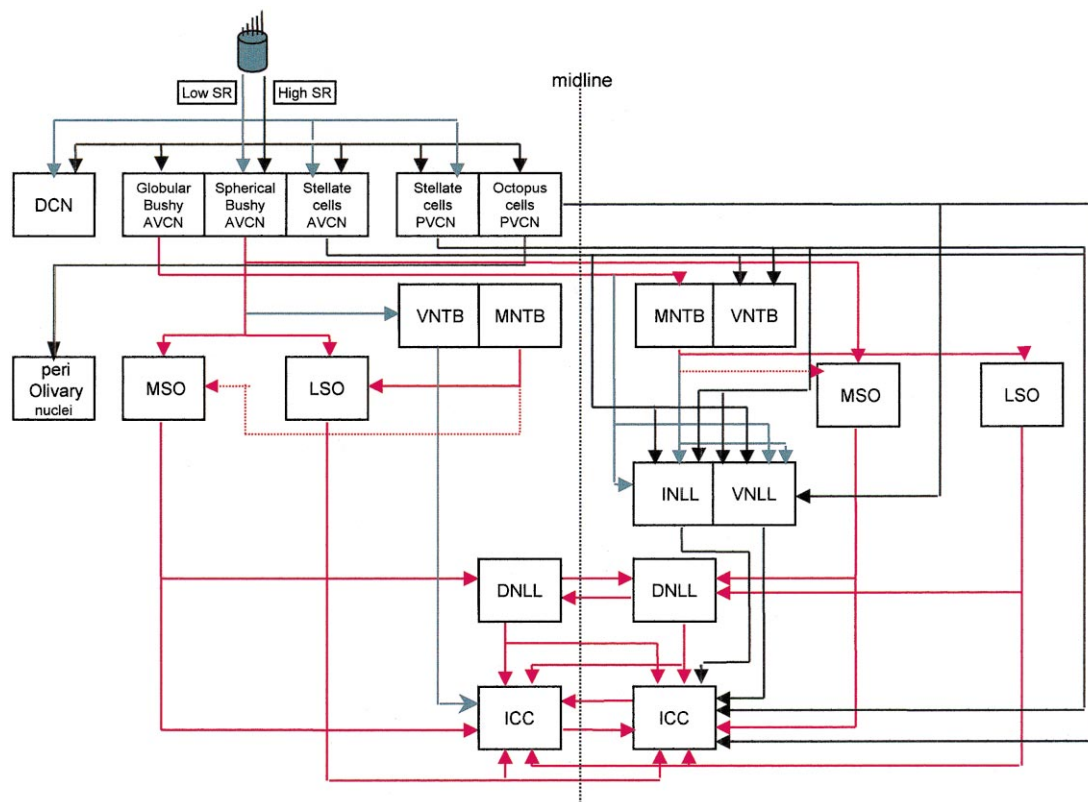


Fig. 2. Simplified scheme of the afferent pathways in the auditory system up to the IC in the cat. Various levels of parallel processing and convergence are noted. First of all, low-SR fibers have partially segregated projections from medium- and high-SR fibers (here indicated by 'high SR'). After the separate processing in the major cell groups of the VCN and DCN the pathways follow a distinct course for the sound localization processes (red) and the sound identification processes (black), or belong to both at some point on their course (gray). All pathways converge in the IC, however, only the ICC is shown. Not all known pathways are included and the output from the DCN is not dealt with at all.

starts in the hair cells and finds its basis in the purely mechanically operated transduction channels (Hudspeth, 1989). The openings and closings of these channels can follow the fastest of sound frequencies so that the hair cell is depolarized and hyperpolarized in complete fidelity and in synchrony with the sound frequency. The synapse between the inner hair cells (IHCs) and the AN fibers is also specialized to reproduce as fine a time resolution as possible (Trussell, 1997; Hudspeth, 1999). This exquisite temporal sensitivity is also the basis for a multiplexed representation of sound: a spectral one and a temporal one. Parallel processing allows the initial segregation of localization and identification pathways.

7.1. Parallel processing between cochlea and IC

The ear performs at least three operations on a complex sound besides localizing it. The first is separating the individual frequency components from several that are simultaneously present (e.g. in vowel formants). This spectral complexity of the stimulus is the main determinant of its perceived timbre. It is also important, together with temporal cues, for determining the pitch

of a complex sound and this information is essential for the auditory system to differentiate between two speakers. The second operation for the ear is to enhance the spectral and temporal contrasts of the resolved frequency components in order to compensate for the poor signal to noise ratios in naturally occurring sounds. The third operation is to extract and abstract the behaviorally meaningful parameters from the results of the peripheral spectral analysis (Plomp, 1976; Evans, 1992). The first two of these tasks are performed in the CN.

The output of the VCN largely follows the anatomical anterior and posterior subdivisions: AVCN spherical cell and GBC output is involved in the localization of sound and projects to the SOC. In contrast, the PVCN is only involved in the identification of sound and its output bypasses the SOC to project to the monoaural nuclei of the LL. The sound localization pathways are indicated in red in Fig. 2. The SBCs in the AVCN preserve and convey the timing information of AN fibers bilaterally to the medial superior olive (MSO). The responses of SBCs to tones are sharply tuned and phase-locked for frequencies below 3–5 kHz. The projection patterns of the large SBCs produce delay lines in

the MSO. Most axons terminate in an isofrequency band in the contralateral MSO with axon collaterals that vary systematically in length, being shortest medially and largest laterally. As a consequence of this projection pattern, neurons in the MSO are activated in a systematic function of the location of sound in the horizontal plane (Smith et al., 1993; Oertel, 1999).

The sound identification pathways in Fig. 2 are indicated by black (completely separate from the localization path) and gray (splitting off the localization path). Stellate cell output forms a major, direct pathway from the CNs to the contralateral midbrain but also to the ipsilateral peri-olivary nuclei and to the contralateral ventral nucleus of the trapezoid body and the ventral nucleus of the LL (VNLL). The sharply tuned, tonic responses of stellate cells to tones carry essential acoustic information: each cell encodes the presence of energy present in a narrow acoustic band with tonic firing. The firing of the population of stellate cells can thus provide a rate representation of the spectral content of sounds.

Almost all of the monaural and binaural pathways from the lower auditory system project to the ICC. Neurons in the ICC project to the auditory forebrain. Thus, the cellular anatomy of the IC may provide the substrate for integration of the ascending pathways on single projection neurons (Oliver and Huerta, 1992). The operation of the IC consists largely of integrating inhibitory and excitatory inputs, which themselves have different temporal properties, and so interact to produce filters for temporal features of sound (Keller and Takahashi, 2000). The filters appear tailored for the analysis of sound envelope; such as estimation of the duration, the envelope modulation rate, or the rate of FM (Covey and Casseday, 1999).

Spatial information represented in IC cells is conveyed both to AI and to the SC in the midbrain of mammals, or to their homologues: field L and the optic tectum in birds. In the midbrain pathway, the location of an auditory stimulus comes to be represented in a map of space. The first step towards creating a space map takes place in non-tonotopic subdivisions of the IC. These subdivisions are sites where information about spatial cues is combined across frequency channels, yielding neurons that are broadly tuned for frequency and tuned for sound source location. Unlike neurons in the tonotopic pathway, neurons in these areas are far more responsive to complex sounds than they are to tonal stimuli. By integrating information across frequency channels in a non-linear fashion, they eliminate spatial ambiguities that are inherent to frequency-specific cues and become tuned for a single sound source location. This transformation, of a frequency-specific code for spatial cues into a topographic code for space in the ICX, has been described in several

species. The output of the space processing regions of the IC is conveyed to the SC where, in all species studied so far, an auditory map of space exists.

At this stage it is appropriate to ask what has happened between AN, the origin of the distributed processing, and the ICC where all the tonotopic pathways converge. In the AN, STRFs describe the preferred sound for individual fibers succinctly and can be used to predict the response to any sound (Kim and Young, 1994). STRFs in the midbrain of frogs (Aertsen and Johannesma, 1980, 1981; Aertsen et al., 1980, 1981; Hermes et al., 1981; Epping and Eggermont, 1985) are more complex than those for the AN. These STRFs only qualitatively predicted the responses of individual cells to complex sounds that were sufficiently different from the Gaussian noise with which the STRF was determined (Eggermont et al., 1983a). In fact, the STRFs in frog midbrain are very similar to those obtained in AI of the cat (deCharms et al., 1998) and field L of birds (Theunissen and Doupe, 1998). Thus one wonders what the predictability for responses to sounds that differ from those that were used to estimate the STRF for these areas will be.

Most of the response types described for the CN are found in the IC. Non-monotonic rate intensity functions appear in the ICC, potentially reflecting similar non-monotonic input from the dorsal CN (DCN) (Aitkin, 1986). The AN and ICC are both tonotopically organized, however, whereas the AN has a smooth representation of CF, the ICC shows a step-wise progression of CF along its main topographic axis thought to reflect a framework for representation of psychophysical critical bands. Combined herewith a smooth frequency gradient exists orthogonal to the main frequency axis (Schreiner and Langner, 1997). Thus the ICC comprises both analytic and integrative properties in the frequency domain. In addition, periodicity information in the cat appears also topographically organized orthogonal to the fine structure of the frequency representation (Langner and Schreiner, 1996). The colocalization, albeit along orthogonal axes, of frequency pitch and periodicity pitch in the ICC may well provide for a similar organization proposed for auditory cortex (Schulze and Langner, 1997).

In order to explain a potential topographic mapping of periodicity in ICC (a periodotopic map) a coincidence detection mechanism between onset activity and chopper activity was proposed (Langner, 1992; Langner and Schreiner, 1996). Combining this with the correlation between best modulation frequency (BMF) and CF the mapping would follow automatically. However, as Krishna and Semple (2000) recently pointed out, because of the dependence of the rate modulation transfer function (rMTF) on sound pressure level (SPL), such a map is likely not level tolerant. In addition,

they considered it as more likely that the band-pass rMTFs seen in IC neurons, interpreted as an indication of rate coding of periodicity, result from coincidence detection of synchronized excitatory inputs. This converts the peak of a temporal modulation transfer function (tMTF) in CN to a rMTF peak in IC. Coincidence detection mechanisms usually destroy temporal information present in phase-locked firings, but if preferred intervals exist they may survive this process (Cariani, 1999b). This loss of temporal information can be seen in the low-pass or band-pass shape of the synchronized firings as a function of modulation frequency in VCN (Møller, 1973). Thus, it is no surprise that the ICC is the first structure where the firing rate is tuned to AM frequency. Consequently, the IC has been proposed as the place where certain aspects of temporal information present in the CNs are transformed into a rate-place code (Langner, 1981). However, because of the periodic nature of the phase-locked firings, the output of the coincidence detector will still be periodic and a representation of synchronized firing still exists in ICC albeit that it is largely identical to that for firing rate (Epping and Eggermont, 1986). Krishna and Semple (2000) in the IC of the cat also demonstrated that most IC neurons responded with significant synchrony to modulation frequencies up to 300 Hz. Thus considerable temporal information remains in the response in the range of modulation frequencies where most BMFs for the rMTFs lie. The emergence of rate tuning in the IC does not necessarily preclude the possibility that information about modulation frequency is also present in a temporal code.

It is tempting to consider the specialization of the auditory system for timing as restricted to the levels below the IC. It is at these levels that accurate temporal representations for space and pitch are found and it is likely that most of these are converted into rate-place representations in the IC. As a consequence, processing of auditory information above the IC should potentially be compared to processing of other sensory modalities. This would allow answering the question if a specialization for processing related to the nature of sound remains. If this is not the case, then neural coding strategies found in the somatosensory and visual modalities may apply also to the auditory system.

7.2. Parallel processing between IC and auditory cortex

The out of the midbrain pathways of the IC (Fig. 3) comprise: (1) a lemniscal pathway (black), that originates in the ICC and terminates in the ventral division of the medial geniculate body (MGBv), and continues to AI. (2) A lemniscal adjunct pathway (fat gray lines), that originates in the cortical (ICX) and paracentral subdivisions of the IC, and targets the deep dorsal,

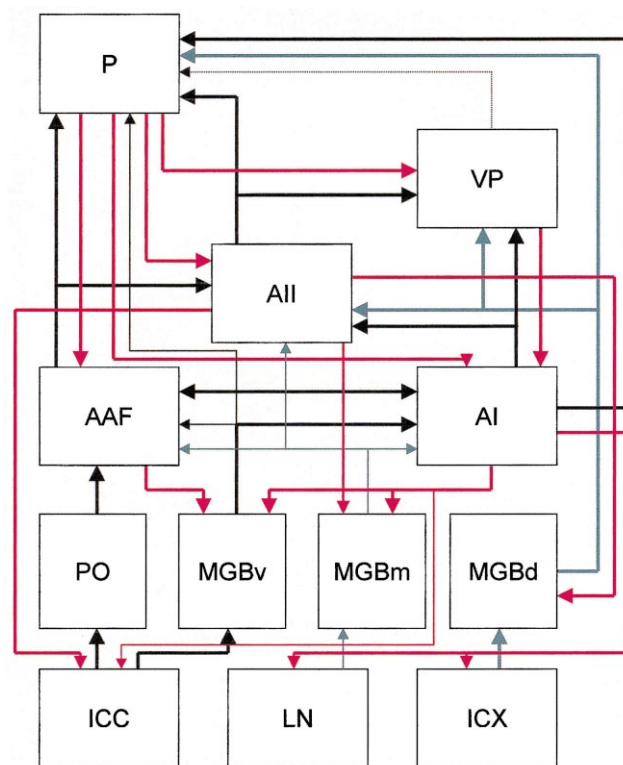


Fig. 3. Simplified scheme of the afferent and efferent connections between the IC and auditory cortex of the cat. Five cortical areas are included and they are portrayed as hierarchical based on Wallace et al. (1991). Most of the connections and their strengths are based on Winer (1990). Exquisite detail about thalamo-cortical connections is provided in Huang and Winer (2000). Distinctions are made between the lemniscal pathway (black) which originates in the ICC, and the lemniscal adjunct pathways (fat gray) which originate in the ICX. In addition there is a diffuse modulating pathway (thin gray lines) that projects to the superficial layer of all known cortical areas. The lemniscal pathway projects to the MGBv and the posterior group of nuclei in the thalamus (PO), and from there to AI, AAF and parts of field P. The lemniscal adjunct pathway projects to AII, P and VP. Strong reciprocal (efferent) connections exist between all cortical areas (fat red lines) and between cortical areas and subdivisions of the thalamus, as well as with the IC but excluding the ICC. Weaker efferent connections exist between AI and ICC.

dorsal and ventro-lateral nuclei of the MGB. This continues to non-primary cortical areas. (3) A diffuse pathway (thin gray lines), that originates from neurons throughout the IC and lateral tegmentum and sends axons to the medial division of the MGB (Aitkin, 1986; Graybiel, 1973; Oliver and Huerta, 1992).

Cortex is a relatively new invention: it is found only in mammals, albeit that a homologue is found in archosaurs (birds and crocodiles). In most other vertebrates the pinnacle of sound processing is a homologue of the IC, such as the torus semicircularis in fish, frogs and toads, and lizards. Obviously, this presents sufficient information processing capacity of sensory events to prevent extinction of these animals.

What does the ‘new’ supra-collicular processing network add to the representation of sound? How different is the sensory processing in this supra-collicular circuit from those in the visual and somatosensory modalities? Integration of sensory information with cognitive aspects is a likely computational outcome and imposes demands that are different from the topographic computational maps found in brainstem and midbrain that provide for fast processing. For instance, auditory space is represented in a clustered format (assemblies) in the forebrain (Cohen and Knudsen, 1999). The forebrain is likely to be essential for the selection of one auditory stimulus out of several possible stimuli, for the identification of sounds, and for remembering the location of stimuli that were heard in the recent past.

Clustering of neurons with similar response properties is a ubiquitous feature of functional organization throughout the CNS. One possibility is that clustered organizations result from competitive interactions among multiple, independent parameters, all competing for representational space in the same area, causing the representation of any single parameter to become severely disrupted. High-order maps (maps of higher order stimulus properties), e.g. those proposed for vowels (Ohl and Scheich, 1997) or phonemes, might be expected in areas of the brain that process aspects of perception or behavior for which the values of a parameter are the essence of the analysis. Examples are rise time or direction and rate of FM in animal vocalizations (Mendelson et al., 1993).

The most salient properties of cortical neurons are: (1) that adjacent neurons generally fire independently under spontaneous conditions (Eggermont, 1992a,b,c) and slightly more synchronized under stimulation (Eggermont, 1994a,b). (2) That cortico-cortical synapses are generally very weak and that the same neuron may be involved in several different processes (Abeles, 1982). (3) That cortical neurons can, under the appropriate conditions, engage in higher processes for a period of over 1 s and patterns of activity across several units have the ability to maintain strict timing even after delays of hundreds of ms. These properties support the view that the computing element in the cortex is a neuron population in which the activity of any given cell has very little impact (Abeles, 1988). An exception to this view has to be made for burst firing; a burst of only two or three spikes in an individual neuron could be sufficient to fire the receiving neuron in some brain areas (Lisman, 1997).

At least seven cortical areas in the cat have a representation of the cochlea and there are potentially seven additional auditory areas (Winer, 1990; Huang and Winer, 2000). For six auditory cortical areas in cat, AI, anterior auditory field (AAF), AII, V, ventral posterior auditory field (VP), and P, information is avail-

able about the mapping of CF on the spatial extent of the area (Reale and Imig, 1980). Only five of those areas are indicated in Fig. 3. Thus, for AI a gradual shift from low to high CF is found from the caudal to the rostral end (Merzenich et al., 1975), that for CFs of about 35 kHz reverses direction, indicating that AAF is reached. At the low-frequency boundary of AI the frequency gradient also reverses by the entrance into field P.

AI in cat is locally very homogeneous: at a given recording site, the CF, threshold at CF, FTC bandwidth and minimum latency of the separated single units were very similar (Eggermont, 1996). AI in cat is not spatially homogeneous, in the dorso-ventral direction, i.e. along the isofrequency sheets the sharpness of tuning is greatest in the medial part and increases towards both boundaries (Schreiner and Mendelson, 1990; Schreiner and Sutter, 1992; Schreiner et al., 2000). This was also found using ripple spectra, broad band signals sinusoidally modulated on a logarithmic frequency scale (Schreiner and Calhoun, 1994; Shamma et al., 1995) that are considered to be the building blocks of all complex sounds (Shamma, 1996; Wang and Shamma, 1995). The maps of characteristic ripple frequency in ferret AI (in cycles/octave, Ω_0) exhibited two trends (Versnel et al., 1995). First, along the isofrequency planes, the largest values were grouped in one or two clusters near the middle of AI, with smaller values found towards the edges. Second, along the tonotopic axis, the maximum Ω_0 in an isofrequency range increased with increasing CF. FTC bandwidth, which was inversely correlated with Ω_0 , exhibited similar distributions. The maps of the characteristic phase (measured in radians relative to the phase of a sinewave starting at the low-frequency edge of the complex, Φ_0) also showed a clustering along the isofrequency axis. At the center of AI symmetrical responses ($\Phi_0 \approx 0$) predominated. Toward the edges, the response fields (RFs) became more asymmetrical with $\Phi_0 < 0$ caudally and $\Phi_0 > 0$ rostrally. The asymmetrical RFs tended to cluster along repeated bands that paralleled the tonotopic axis. The FM directional sensitivity (DS) tended to have similar trends along the isofrequency axis as Φ_0 . The findings suggest that AI cells can, in principle, function as ripple band-pass filters, analyzing an input spectral profile into separate channels tuned around different characteristic ripple frequencies. Equivalently, from the perspective of their response area bandwidths, they can be said to have a range of bandwidths so as to analyze the input spectral profile into different scales (Shamma, 1996).

He et al. (1997) have suggested that the extreme dorsal part of AI forms a separate cortical area, specialized by its capacity for temporal integration and tuning to stimulus duration. This area also features long latencies

and broad multi-peaked FTCs (He and Hashikawa, 1998; Sutter and Schreiner, 1991). The FTCs in the dorsal part had a lateral inhibitory structure that was distinct of that in ventral AI: specifically lateral suppression areas flanking the excitatory tuning curve on both sides were much less common (Sutter et al., 1999; Versnel et al., 1995). On this basis, Sutter et al. (1999) suggested that the dorsal part of AI is involved in analyzing complex spectra, whereas the ventral part would be poorly responsive to such broad band stimuli. In AAF, the high frequencies are located dorso-caudally and lower frequencies are found into the rostro-ventral direction (Knight, 1977; Phillips and Irvine, 1982).

AII in cat appeared not as well organized tonotopically as AI and the units showed broader frequency tuning and higher thresholds to tone burst stimulation than in AI (Schreiner and Cynader, 1984). Recently, a double representation of the cochlea within the classical boundaries of AII in cat was demonstrated (Volkov and Galazyuk, 1991). One tonotopic representation was found in the dorso-caudal region (2.6–3.2 mm long) with a spatial orientation similar to that described in AI (low frequencies caudal, high frequencies rostral), the second ventro-rostral region was smaller (1.4–2.5 mm) and had the opposite tonotopic orientation (i.e. as in AAF). The dorso-caudal region may be close to the transition zone between AI and AII as defined by Schreiner and Cynader (1984). Alternatively it could be the ventral auditory cortex (Ve) as defined by Huang and Winer (2000). Thus AII, not unlike AI, may consist of several specialized subregions.

The posterior field of cat auditory cortex (P) is also characterized by a tonotopic organization, narrow V-shaped tuning curves and a dominance of non-monotonic rate intensity functions for tone pips with short rise times. The latencies are generally longer than in AI, the accuracy of first-spike latencies is likewise poorer (Heil and Irvine, 1998b). Thresholds are invariably higher than in AI (Reale and Imig, 1980; Phillips and Orman, 1984; Phillips et al., 1995). Field VP, caudal from field P, also has a tonotopic organization as has field V found caudally from AII and bordering the posterior ectosylvian sulcus (Reale and Imig, 1980). Non-tonotopically organized areas are field DP dorsal from AI and bordering the anterior ectosylvian gyrus, field T rostro-medial from AII (He and Hashikawa, 1998; Reale and Imig, 1980), and field AIII medial from AAF (Winer, 1990).

At supra-threshold levels, the well organized topographic map in AI broadens in the direction of higher CFs and partially breaks down along the isofrequency contour to acquire a patchy appearance (Phillips et al., 1994). This has been attributed to the presence of alternating aggregations of monotonic and non-monotonic

units along the isofrequency contours. Heil and Irvine (1998a) show that in cat AI there are orderly topographic organizations, along the isofrequency axis, of several neuronal properties related to the coding of the intensity of tones. These are minimum threshold, dynamic range, best SPL, and non-monotonicity of spike count – intensity functions to tones of CF. Minimum threshold, dynamic range, and best SPL are correlated and alter periodically along isofrequency strips (see also Schreiner et al., 2000). The steepness of the high intensity descending slope of spike count–intensity functions also varies systematically, with steepest slopes occurring in the regions along the isofrequency strip where low thresholds, narrow dynamic range and low best SPLs are found. As a consequence, CF tones of various intensities are represented by orderly and, for most intensities, periodic, spatial patterns of distributed neuronal activity along an isofrequency strip. For low to moderate intensities, the mean relative activity along an entire isofrequency strip increases rapidly with intensity, with the spatial pattern of activity remaining quite constant along the strip. At higher intensities, however, the mean relative activity along the strip remains fairly constant with changes in intensity, but the spatial patterns change markedly. As a consequence of these effects, low and high intensity tones are represented by complementary distributions of activity alternating along an isofrequency strip. It was concluded that in AI tone intensity is represented by two complementary modes, viz. discharge rate and place. Besides that, sharpness of tuning, response strength (combination of threshold and dynamic range) and temporal changes of stimulus spectrum (FM rate and DS) show independent topographical organizations within the isofrequency contours and this suggests parallel and independent processing of these acoustic aspects in AI (Heil et al., 1992b).

Thalamo-cortical and intrinsic cortical connection patterns of AI in cat indicate a modular organization (Wallace et al., 1991; Clarke et al., 1993; Schreiner et al., 2000; Huang and Winer, 2000). Single thalamo-cortical fibers branch and terminate in patches in layers II–VI. Individual neurons in MGBv can be double-labeled when two tracers are injected at different dorso-ventral positions along the isofrequency axis. Thus, spatially segregated subsets of neurons along the isofrequency axis of AI share a common input from the thalamus. It is possible that subregions of AI that receive common thalamic input also mutually innervate each other via horizontal connections in layer III. Thalamo-cortical patches in rabbits and cats are segregated by approximately 1.5 mm. This agrees with the finding that neuron pairs 0.2–1.0 mm apart tend to have less interactions than pairs 1.0–2.0 mm apart (Clarke et al., 1993; Eggermont, 1993b). Adjacent intrinsic clusters in the

narrowly tuned region of AI in cats are segregated at spatial intervals of the same range. Long-range intrinsic cortical connections in AI occur between clusters of neurons with similar CFs in an elongated patchy pattern that follows, and is confined to, the dorso–ventral isofrequency axis. A fundamentally different and more complex pattern, including a large degree of frequency convergence from patches outside the isofrequency axis, was observed with injections into the broad band subregions of dorsal and ventral AI. The data provide anatomical evidence for at least two spatially segregated systems of spectral integration in the AI that may be related to the distinction between CB and non-CB integration behavior.

In behaving macaque monkeys, a tonotopic organization exists for the AI but not for the surrounding (belt and parabelt) areas. Topographic organization of other response parameters, previously demonstrated in the anesthetized cat, was not apparent in the behaving monkey (Recanzone et al., 2000a). Some evidence for a hierarchical processing of neural activity related to sound localization, proposed by Rauschecker (1998) as part of a ‘where’ and ‘what’ pathway segregation (Rauschecker and Tian, 2000), was found between AI and area CM in behaving macaques (Recanzone et al., 2000b; Romanski et al., 1999). In general, whereas in cat most cortical fields are directly innervated by divisions of the MGB (Huang and Winer, 2000), the functional connection pattern in primates appears to be more hierarchical. The input from the ventral division of the MGB activates the three core areas of auditory cortex, these activate the numerous belt areas, which in turn activate the parabelt areas (Kaas and Hackett, 2000). Although there are direct projections from dorsal and medial MGB to the belt areas in primates, these appear to modulate rather than activate. The almost complete interconnectivity pattern of auditory cortical areas in cat is not present in primates; there are no direct connections from the core to the parabelt areas (Kaas and Hackett, 2000).

7.3. Efferent connections introduce non-linear dynamics into the auditory system

The auditory system is not just an afferent projection system but has a myriad of efferent connections that makes it a reentrant system characterized by multiple, loosely interconnected, regional feedback loops (Spangler and Warr, 1991). At the lowest level, a loop between cochlea and the SOC comprising the olivocochlear bundle exists. A second loop is found between the lower brainstem nuclei and the IC. A third loop is formed between the IC and the thalamo–cortical system which, in itself, consists of a feedback loop between thalamus and cortex. In Fig. 3, efferent connections

are shown in red. More specifically, the auditory cortex projects back to the MGB with 10 times as many fibers than the number of afferents from the MGB to auditory cortex. The auditory cortex also connects with the IC, but with exclusion of the central nucleus (Winer, 1990). The central and external IC subnuclei both project back to the DCN. The DCN in turn feeds back to the VCNs. It seems that the strongest contiguous projections from cortex to the periphery involve the nuclei of the extralemiscal pathway, including the DCN.

The olivocochlear bundle projects via its medial branch to the outer hair cells, thus regulating the slow motility of the outer hair cells and thereby the stiffness of the basilar membrane. Via its lateral branch it especially affects the low spontaneously active AN fibers synapsing with the IHCs (Spangler and Warr, 1991; Guinan, 1996). Activation of the olivocochlear bundle appears to improve the discriminability of signals, such as speech, in the presence of broad band noise by lowering the noise-induced background firing rate and increasing the dynamic range (Winslow and Sachs, 1987; Liberman, 1988). Chronic cochlear de-efferentation in adult chinchillas resulted in reduced spontaneous AN fiber firing rates, increased driven discharge rates, decreased dynamic range, increased onset to steady-state discharge rate, and hypersensitive tails of the frequency tuning curves (Zheng et al., 1999). Clearly the efferent system plays a large role in maintaining the normal operating mode of the cochlea.

The role of this nested set of reentrant systems can only be speculated upon. One possible role is an involvement in expectancy. A subject has some expectancy about the probability of various environmental occurrences stored in its internal representation and incoming sensory information is continuously checked whether it corresponds, within certain error, to this internal representation. This requires reentrant, creative, systems (Edelman, 1987; Johannesma et al., 1986) with a definite role for thalamo–cortical and cortico–cortical loops. In fact, Mumford (1991, 1992) proposed that the thalamus acts like ‘an elaborate seventh layer of the cortex’ and functions as a ‘blackboard’ on which the cortex writes and from which it reads. This could be the place where the internal representation of the world is stored and continuously updated. Support from this notion comes from studying the effect of deactivating AI on neural interaction in the thalamus (Villa et al., 1999). A major effect appears to be, at least in the cat, a disinhibition of inhibitory interneurons as a result of inactivation of projections from the reticular thalamic nucleus.

The effect of cortical activity on the auditory thalamus was first demonstrated by cooling the surface of AI (Villa et al., 1991) resulting in a decrease of spontaneous activity and changes in the CF of the frequency

tuning curve by up to 0.5 octave. Recently, the role of the corticofugal systems has received new attention especially through the work of Suga and collaborators. They found that electrical stimulation of delay-tuned neurons in the mustached bat auditory cortex increased the responses of those neurons in the IC that were tuned to the same echo delay; conversely it decreased the responses to neurons tuned to different echo delays (Yan and Suga, 1996, 1999). This modulatory action of the cortex is potentially involved in the adjustment to long-term changes in the overall functional organization of the IC, MGB and AC. Subsequently, corticofugal activity was also shown to be involved in subcortical frequency tuning in other bats as well as in cats. This was likely caused by positive feedback which, in combination with widespread lateral inhibition, sharpens and adjusts the tuning of neurons in the IC and thalamus (He, 1997; Jen et al., 1998; Zhang et al., 1997; Yan and Suga, 1998; Zhang and Suga, 2000; Suga et al., 2000).

The presence of the feedback loops thus creates a non-linear dynamic system. A crucial property of non-linear dynamic systems is the exponentially increasing amount of uncertainty associated with predicting the future time evolution of the system and its related information content. This is determined by the sensitive dependence of the system upon the initial conditions so that it takes only a very small perturbation to switch the system from one unstable state into another. Any stimulation could act as such a disturbance. But what does this have to do with neural coding? First of all we have to accept that brain activity can be modeled as resulting from a non-linear dynamic system. If we assume that behavior is the only relevant correlate of brain activity, then the brain may be considered to operate at the edge of chaos. This is a state where very small changes in the controlling parameters (such as sensory stimuli) may cause large macroscopic state changes such as from one quasi-periodic state into another (Schiff et al., 1994). Alternatively, the brain may be a stochastic system and still be sensitive to small perturbational feedback just as in a deterministic chaotic system (Christini and Collins, 1995). Under these constraints, coding by individual neurons is largely irrelevant, i.e. it does not matter whether a single-unit rate code or a temporal code is at work. The only relevance of neural firing is to be found in its capacity to induce transitions between the many unstable macrostates found in the brain and the relevant questions are related to what determines the various transition probabilities. Thus only population activity is relevant and specifically the spatio-temporal patterns that determine the transitions between states are the relevant ones to investigate. Any change in the firing patterns of neurons that causes a transition between macrostates, for in-

stance a single-unit burst or an intricate multi-neuron activity pattern, now becomes important.

State switching in brain activity may thus be more relevant than the neuronal patterns and the correlations between them that are found during stable states. State switching is likely reflected in changes in synchrony between individual neuron activities (Vaadia et al., 1995). Time-dependent changes in neural synchrony (Aertsen et al., 1989) may signal the building up or breaking down of assemblies. When control parameters (e.g. concentration of neurotransmitters, drugs, etc.) change beyond a critical value, the system suddenly forms a new macroscopic state that is quite different from the previous one. Receptor systems in the brain are organized to detect changes in transmitter concentration as opposed to absolute concentration (Black, 1991) which stresses the notion that change has higher behavioral value than steady states. Thus, the neural correlates of transitions between brain states may reveal more of neural coding than an exhaustive description of single-unit properties in relation to these brain states. Relational order, feature binding and the emergence of wholes (Gestalts) as revealed by synchronous population neural activity, may supersede the single- or few-unit activity description of brain function (Merzenich and deCharms, 1996). Relational order between individual neurons is equivalent to the existence and strength of correlation maps (Koenderink, 1984) and sees assemblies as more than a statistical result of the outcome of a large number of interactions among discrete neurons. Through the reciprocity of connections between the neurons in the map, which may provide a positive feedback, very small changes in just a few pair correlations may give rise to a sudden state change characterized by the formation of a new assembly. That only very small changes are required to induce the transitions between states is suggested by the sudden increase in intercortical area synchronization between silence and low level acoustic stimulation (Eggermont, 2000a). Are such state transitions operative when perceiving speech? Or are they only operative when switching attention to sounds signifying danger or potential food? It appears that paying attention substantially increases neural synchrony between neurons in somatosensory cortex (Steinmetz et al., 2000). Before I continue with neural coding pertinent to sound perception, I will first review the neural representation of speech features in the auditory nervous system.

8. Representation of selected speech sound features across the auditory pathway

I will describe the neural processing of speech sounds by initially focussing on some characteristic elements of

speech such as vowels, fricatives and consonants. Then I will describe the representation of formant transitions and FM, VOTs and AM. This will be followed by a survey of the neural representation of vocalizations. For all of this one has to keep in mind that standard speech and other species-specific vocalizations are usually produced, and received, at a SPL at which most high-spontaneous firing rate (SR) AN fibers are already saturated in their firing rates to steady-state sound. Secondly, speech perception can tolerate background noise levels that are only 10 dB lower than the speech. Thus any neural representation that is a candidate for transformation into a neural code should be tolerant of these conditions.

8.1. Representation of sound spectrum

Starting with the first classic studies of vowel representation in the AN (Sachs and Young, 1979; Young and Sachs, 1979) the tone was set for contrasting representations based on average firing rate and those based on phase-locked firings and thus focusing on periodicity information. Yet both representations were topographic: firing rate or firing intervals were evaluated as a function of CF for the population of AN fibers.

8.1.1. Vowels

Sachs and Young (1979) measured the discharge rate of single AN fibers in cats to steady-state vowel stimuli: /I/, /e/ and /a/ with a 128 Hz fundamental and the formants placed on multiples of the fundamental frequency. As expected, the profiles of discharge rate vs. fiber CF showed peaks which corresponded to the formant energy peaks in the vowels when the stimuli were presented at low sound levels. At higher levels, generally those corresponding to normal speech levels, the peaks in the rate profiles disappeared, principally because of rate saturation. Only the small group of low-SR fibers (which constitutes less than 5% of the type I afferents) was capable of representing the spectral peaks in the vowel spectrum up to high SPL.

Young and Sachs (1979) computed the average localized synchronized rate (ALSR), the average value of the response at each harmonic (of the fundamental period of 7.81 ms), expressed as a synchronized firing rate, and calculated over all fibers whose CFs were within 0.5 octave of the frequency of the harmonic. At low stimulus levels, synchronized responses were usually observed at the frequency of each harmonic in the stimulus; and were maximal in fibers with CFs corresponding to their frequency. Larger amounts of synchrony were observed for the larger stimulus components at the formant frequencies. As sound level is increased, the response to the first formant grows in

amplitude and spreads out through the population of AN fibers, but most rapidly in the direction of higher CFs. At the highest level used (84 dB SPL) the temporal responses were almost entirely dominated by responses to the first formant.

In constructing the ALSR functions it was assumed that the CNS could perform a limited spectral analysis of the responses of auditory neurons: it is only necessary to extract one frequency component from the response of each fiber. Whether the CNS is capable of such a calculation can be questioned, but the fact that central neurons are sensitive to the small ITDs necessary for binaural hearing suggests that calculation of the kind needed is at least feasible. Palmer et al. (1986) replicated the findings in guinea pig AN for the vowels /i/, /a/, and /u/. They also found that primary-like cells in VCN showed less synchronization for the second and third formant frequencies at least for the /i/ and /u/.

Profiles of average discharge rates versus CF (rate-place representation) can show peaks in the vicinity of formant frequencies when the vowel /e/ is presented alone. These profiles change drastically in the presence of background noise, however. At moderate vowel and noise levels and S/N of +9 dB, there are no peaks of firing rate near the second and third formant frequencies (Sachs et al., 1983). When rate profiles are plotted for low-SR fibers, formant-related peaks appear at stimulus levels higher than those at which peaks disappear for high spontaneous fibers. In the presence of background noise, however, the low spontaneous fibers do not preserve formant peaks better than do high spontaneous fibers. Representations that reflect phase-locked properties as well as discharge rate (temporal-place representations) are much less affected by background noise. Plots of ALSR versus CF show clear first and second formant peaks at all vowel and noise levels used. Except at the highest level (vowel 85 dB SPL, S/N = +9 dB) there is also a clear third formant peak. This pattern was also found in cats exposed to acoustic trauma: there is essentially no rate information, which would allow the vowels with identical *F1* but different *F2* frequencies to be discriminated. However, information that could allow discrimination remains in the temporal aspects of the responses (Miller et al., 1999).

Delgutte and Kiang (1984a) confirmed Young and Sachs' findings for a set of nine vowels and noted that all the available physiological data show that responses of most auditory nerve fibers to vowels are dominated by components near the formant frequencies. In contrast, responses to broad band noise stimuli are dominated by components near the fiber CF, comparable to those that are extracted using reverse correlation (de Boer and de Jongh, 1978). This contrast in the distribution of response components between broad

band noise and vowel-like sounds may help the central processor in distinguishing speech from other environmental sounds, although dynamic cues would also play a role in continuous speech.

These early studies clearly favored temporal representation of vowels, but a clear mechanism in the CNS that could use such a representation could not be identified. This prompted a renewed search for evidence that rate–place representations were capable of discriminating slightly different vowels. By analyzing the differences in neural firing rates to pairs of stimuli with the same $F1$ but slightly different $F2$ frequencies, Conley and Keilson (1995) showed a robust rate representation. Second formants differing by 125–140 Hz could be discriminated using individual fibers optimally placed on the basilar membrane. For the whole AN the just noticeable difference for $F2$ was estimated at 1 Hz. May et al. (1996) and Le Prell et al. (1996) subsequently showed that the discharge rate of fibers with best frequencies near $F2$ provides enough information to allow discrimination of these stimuli at the performance levels shown by cats in behavioral experiments. They used a signal detection theory analysis to show that this rate difference cue predicts behavioral performance that meets or exceeds the abilities of normal hearing cats.

The distribution of the relative phases of synchronized activity emerges as an important response feature reflecting the stimulus spectral parameters. These relative phase changes between different frequency components occur only at or very near the points of resonance in the cochlea. Here the responses exhibit drastic relative phase shifts as a result of the sudden slowing down of the traveling wave and the consequent rapid accumulation of phase shifts. These rapid phase shifts thus manifest themselves as steep and localized spatial discontinuities in an otherwise relatively uniform instantaneous pattern of activity across the fiber array, all occurring at the CF locations corresponding to the low-order harmonics of the stimulus (Shamma, 1985a). Response phase is essentially a temporal cue, however, it has also a spatial manifestation because phase shifts along the time axis are accompanied by changes along the spatial, frequency, axis. This dual nature of the phase cue allows its extraction from the spatio-temporal patterns in AN fibers to be reduced to a spatial task. The temporal periodicities of the nerve responses can be used as spatial markers to the CFs of the stimulus components, rather than as absolute measures of their frequencies. This is a fundamentally different approach to the use of temporal cues since no attempt is made here to measure or utilize the actual frequency of locking, but rather only the relative phase shifts of the locked waveforms. Shamma (1985b) showed that at low levels of intensity, a lateral inhibitory network (LIN) can sharpen the average rate profiles. At moder-

ate and high levels the LIN uses the cues available in the distribution of phases of the synchronous activity. These temporal phase shifts manifest themselves at the input of the LIN as steep and localized spatial discontinuities in the instantaneous pattern of activity across the fiber array. The LIN enhances its output from these spatially steep input regions while suppressing its output from spatially smooth input regions where little phase shift occurs. Such a network could be implemented as early as the VCN, where lateral inhibitory activity is present in the type II and type III cells (Rhode, 1991).

Because stellate cells in the AVCN were capable of representing vowel spectra across an extended dynamic range (Blackburn and Sachs, 1990), the rate–place representation of vowel spectra was back at center stage. Recently more confirmation was presented for AVCN and PVCN in the chinchilla (Reccio and Rhode, 2000). They showed that primary-like neurons provided a robust temporal–place representation analogous to the one found by Winter and Palmer (1990) in the guinea pig. In contrast, the chopper cells showed a level tolerant rate–place representation as recently also emphasized for the cat VCN (Keilson et al., 1997). The rate–place coding properties of chopper cells were confirmed and extended by May et al. (1998) who noted that the vowel encoding properties of VCN chopper cells support previous interpretations that patterns of AN convergence on CN neurons compensate for the limitations in the dynamic range of peripheral neurons. Young (1998) noticed that chopper neurons receive both high- and low-SR AN fibers and could respond to high-SR input at low sound levels and low-SR inputs at high levels and so select the best of two worlds. This could be done if high-SR terminals are positioned in the distal portion of the dendritic tree and intermingled with strong inhibitory inputs, the effect of the high-SR inputs can then be cancelled by activation of the strong inhibitory inputs. If low-SR fiber terminals are on the proximal dendritic tree, their effect on the cell would not be affected by the inhibitory inputs. Direct evidence for this idea, however, is lacking. Thus chopper neurons could provide a representation of the shape of the stimulus frequency spectrum which is stable as level changes. This representation could provide information about the identity of the sound: e.g. one speech sound versus another.

Representation of vowels would at least need to represent the two dominant formants $F1$ and $F2$, resulting in a need for combination sensitive units to represent this. The problem is that there are no two orthogonal frequency axes in auditory cortex. Ohl and Scheich (1997), however, suggest that an analogy with experiments using ripple spectra, that demonstrate a systematic representation of the periodicity of the modulated

envelope on the cortex (Kowalski et al., 1996; Schreiner and Calhoun, 1994; Shamma et al., 1995), can be used. Versnel and Shamma (1998) examined whether the vowel responses could be predicted using a linear ripple analysis method, i.e. by cross-correlating the RF of the single unit (obtained from an inverse fast Fourier transform of the ripple transfer function) with the smoothed spectral envelope of the vowel. In most (76%) of AI cells, the relative responses to the natural vowels could be predicted on the basis of this method. In other words, the vowel responses are based on a linear summation of responses to the ripple components of the smoothed spectral envelope of the vowel.

The ripple pattern can be interpreted as a neuronal sensitivity for interactions between spectral components. Ohl and Scheich (1997) suggest to use a transformation of the $F1$, $F2$ map into an $F1$, $F2-F1$ map which leaves the gross structure of the map unaltered. They also suggested that $F2-F1$ is a potentially useful parameter for vowel separation. In experiments with metabolic imaging and electrophysiological recordings in gerbil AI they demonstrate that a formant sensitivity along the tonotopic axis, combined with an $F2-F1$ interaction within the isofrequency contours, conserves vowel separation. In the imaging studies, vowel-evoked patterns in AI appeared as stripes of increased metabolic activity parallel to the dorso-ventral isofrequency contours. The position of the stripe was vowel-dependent: the dorsal boundary of the activated region was located more ventrally as the formant distance $F2-F1$ decreased in the sequence /i/, /e/, /u/, and /o/. Tests with single formants $F1$ or $F2$ produced stripes of labeling with maximal dorsal extent. In the rostro-caudal direction stripes were centered to tonotopic positions corresponding to frequencies between 1 and 2 kHz, but did not allow differentiation between vowels. This indicated that inhibition plays a role in the vowel-specific reduction of neuronal activity along the isofrequency axis. This was confirmed by electrophysiological study of the interaction using synthetic vowels with fixed $F1$ and slowly changing $F2$. Thus, vowels are potentially represented in auditory cortex by activity distributed in the low-frequency areas along the tonotopic axis, without resolving formant peaks, but by mapping relations between spectral peaks in a topographic manner along the isofrequency axis. Consequently, the ultimate result of all the transformations along the auditory pathway could be an orderly map of vowels based on formant separation. It would be interesting to explore the level tolerance of this representation in light of the disintegration of the tonotopic map in AI for higher intensity levels (Phillips et al., 1994).

8.1.2. Consonant-vowels and fricatives

Extensions of the study on vowels were made by us-

ing consonant vowels (CVs) (Sinex and Geisler, 1983; Miller and Sachs, 1984; Delgutte and Kiang, 1984c; Carney and Geisler, 1986; Deng and Geisler, 1987) and fricatives (Delgutte and Kiang, 1984b). The same general pattern emerged: 'phonetic distinctions among many steady-state voiceless fricatives are based more on the general location of the bands of energy in the stimulus spectra than on the existence of spectral peaks associated with formant frequencies. Thus a sufficient requirement for response measures based on AN fiber data is to represent the gross shape of the spectra of fricative stimuli' (Delgutte and Kiang, 1984b). One could conclude that average discharge rates provide more information about the spectra of formant transitions than they do about the spectra of steady-state vowels. The synchrony of discharges, however, may provide more detailed information about the spectra of CVs, than does average discharge rate. Each fiber's response may be characterized by the dominant response component, the largest peak in the Fourier transform of the period histogram. The trajectories of the first three formants can be inferred from changes with time in the dominant components in a sample of fibers (Sinex and Geisler, 1983). When temporal properties of the neural responses, i.e. as reflected in the PSTH, and stimulus spectra are displayed in spectrographic format it becomes clear that especially high-SR nerve fibers synchronize to formants throughout the stimulus and faithfully reflect the formant transitions (Carney and Geisler, 1986).

Both rate-place or temporal pattern-place representations emphasize the preservation or enhancement of the spectrogram of the sound. The population of AN fibers can be seen as a filter bank and the output thereof, in terms of number of spikes in a certain integration window, signals the relative level of the spectral peaks and tracks the changes of the formants as a function of time. The temporal pattern-place representation needs an additional calculation to allow the use of the synchronized rate at each narrow band filter. These local filtering schemes can be related to the concept of 'central spectrum'. A central spectrum is the result of running a spectral analysis of acoustic stimuli in the peripheral stages of the auditory system. Estimates of the frequency selectivity of this analysis for the 2 kHz region in the human are usually between 1/6 and 1/3 octaves. According to this view, the perception of the pitch of complex tones and vowel quality would essentially be pattern recognition operations on the central spectrum. Because AN fibers are frequency selective, the central spectrum could in principle be obtained from the profile of average discharge rates against CF (Delgutte, 1984).

It is known that the discrimination of speech by human listeners is not severely degraded by background

noise at S/N down to 0 dB. We have seen that the temporal representation of vowels under noise is very robust. What about the consonants? Spectral characteristics of both consonant and vowel segments of the CV syllables appear to be strongly encoded by, high-SR, AN fibers in the cat at S/N ratios of 30 and 20 dB. At S/N=10 dB, formant information during the vowel segments is all that was reliably detectable in most cases. Even at S/N=0 dB, most vowel segments were detectable, but only with relatively long analysis windows. The rate responses to the ‘release’ and to the voicing of the stop consonant syllables were quite robust, being detectable at least half of the time, even at the highest noise level (Geisler and Gamble, 1989). In addition, lower-SR fibers (SR < 20 sp/s) encoded important information about CV syllables in the temporal domain more robustly in noisy situations than did high-SR fibers. At low noise levels (S/N=30 dB) even the rate profiles of the lower-SR fibers gave a clearer reflection of a syllables energy spectrum than did those of high-SR fibers (Silkes and Geisler, 1991).

As Geisler (1988) has emphasized ‘the response patterns of the AN fibers still appear to encode only the actual stimulus waveforms’. There does not seem to be any sorting of stimuli into ‘categories’ such as ‘voiced’ or ‘unvoiced’. The onset of voicing, for example, is accurately registered by the enhanced discharge probabilities of low-frequency nerve fibers, whether it occurs early or late in the syllable. Extraction of the phonemic characteristics of the input signals must therefore take place more centrally. It is unlikely that this occurs in the CN.

8.2. Representation of VOT and other gaps in sound

Speech perception not only depends on spectral differences but also on differences in the temporal structure of sound. A voiced stop consonant such as in /b/ can often be distinguished from its voiceless counterpart in /p/ by temporal cues. For stops in the word-initial position, such as in /ba/ and /pa/, the voicing distinction is generally based on the length of the interval between consonant release and the onset of voicing (cf. Fig. 1), this interval is termed the VOT. The perception of a continuum of speech sounds differing only in VOT is categorical (Liberman et al., 1967). In English speaking adults, chinchillas and monkeys, the category boundary between /ba/ and /pa/ is around 30 ms (Morse and Snowdon, 1975; Kuhl and Miller, 1978). When the consonant pair /d/ and /t/ occurs in the word-final position, such as in /had/ and /hat/, the length of the interval between consonant closure and consonant release may distinguish between voiced and voiceless stops; this is referred to as the consonant closure interval (CCI). CCIs have been reported to be off the order of 70–90

ms for voiced stops and 98–130 ms for voiceless stops. For stops in word-medial position, such as in /ada/ and /ata/, /d/ is distinguished from /t/ by the length of the combined interval from closure onset to release and from release to the onset of voicing in the vowel that follows the consonant (Sinex and Narayan, 1994).

For the word-initial stop consonants, peak AN fiber discharge rate and peak response latencies elicited by the syllables with the shortest, /d/, and longest, /t/, VOTs were highly variable across groups of neurons with similar CFs. For VOTs from the middle of the continuum, peak responses were larger, and response latencies were nearly constant across the same group of neurons. The degree of overlap between the discharge rate recovery functions for the 10 and 20 ms VOTs, and the 60 and 70 ms VOTs was extensive, even during the times when stimulus differences were presumably greatest. In contrast, responses to the stimulus pair with VOTs of 30 and 40 ms were more distinctive. For these sounds, the onset of voicing elicited steep discharge rate increases that preserved information about the time at which voicing began. Overall, there was little overlap between these recovery functions, consistent with the finding that psychophysical acuity is usually greater for these syllables. There was no difference in using synchronized discharge rate over average discharge rate (Sinex and McDonald, 1989). There was also extensive overlap of peak response latencies for syllables within the VOT range normally associated with the consonant /d/ (VOT=0–30 ms) and within the range associated with /t/ (VOT=40–80 ms). However, by this measure there was no overlap across categories (Sinex et al., 1991).

For stops in the medial position, analysis of the latencies of the discharge rate changes revealed at least two potentially important patterns of variance. First, the closure onset latencies of neurons with similar CFs could differ by 50 ms. Second, the mean closure onset latency varied with CF, with higher CF neurons signaling onset of closure as much as 50 ms earlier than low CF neurons. Because of this variability in the representation of temporal features of the stimuli, the ability of the CNS to measure the duration of the consonants would be limited. The constraints imposed by the variability in the neural latencies tended to obscure differences between utterances with the same medial consonant, while preserving the contrast between utterances with different consonants (Sinex and Narayan, 1994).

Many aspects of the patterns of response elicited by VOT syllables in ICC (Chen et al., 1996) resembled the patterns elicited from primary auditory neurons with the same syllables. The similarity of discharge patterns to those of primary auditory neurons arises because neurons in both locations accurately reflect the details

of the stimulus spectrum compared to the frequency tuning curve of the neuron. Except for the same reduced variability in the neural coding of VOT in the 20–50 ms range as found in AN fibers, no indication of a categorical perception boundary was found.

Sinex and colleagues stress the loss of temporal information in the neural representation of stop consonants. This loss is greatest in the within category conditions and less across category boundaries. This loss of information produces regions of low acuity and thus may contribute to the formation of a kind of invariant neural code for the temporal properties of stop consonants.

Whatever the variability in the periphery and auditory midbrain imposes on the response in auditory cortex, the response there is typical of the double onset type; an onset response to the initial noise burst and a second-on response at the start of the vowel (Steinschneider et al., 1982, 1990, 1994). The presence of either on response appears highly dependent on the stimulus level. For low CF neurons, intensity levels above 45 dB SPL for a /ba/–/pa/ continuum, typically produced a strong onset response to the noise burst, levels below 65 dB SPL also typically produced a second-on response. As a consequence a double-on response was only clear in the 45–65 dB SPL range, coincidentally close to the level of natural speech (Eggermont, 1995a). In addition the range of VOTs that produced a double-on response was restricted to 30 ms and above. Exceptions could be found at lower SPL where a weak on response to the noise burst could be followed for all VOT values (in a 0–70 ms range) by an equally weak second-on response. So there is no level tolerant coding of VOT for the /ba/–/pa/ continuum in single cortical neurons. In addition there is no natural VOT boundary that distinguishes /ba/ from /pa/ in individual cortical neurons, except in the 45–65 dB SPL range.

Are VOTs detected by intrinsic neuronal properties, as suggested by modeling studies (Eggermont, 2000b) or by neural network properties as advocated by Buonomano and Merzenich (1995)? Both models assume that the neural activity bracketing the gap converges on the same neuron or input channel of the neural network. In the first model, the combination of synaptic depression and amount of after-hyperpolarization can read a range of intervals, and the model was successful in explaining the dependence on the duration of the leading burst. In the latter model, a network of neurons with fixed synaptic facilitation and slow inhibitory post-synaptic potentials, different time intervals are represented by using a population code, and based on the relative activity of different units. This is a transformation from a temporal representation into a place representation. This place representation may be local and consists of clusters of neurons ‘tuned’ to different classes of VOTs, or

the place representation may be global and consists of a topographic map. The advantage of a neural network model is its flexibility and adaptability to VOTs occurring in particular languages. An extensive study on gap and VOT detection properties of neurons in AI, AAF and AII of the cat (Eggermont, 1999) provided evidence for a widely distributed range of gap detection thresholds. Neurons recorded on the same electrode generally had similar gap thresholds, i.e. no evidence for sensitivity to small (< 30 ms) and large (> 50 ms) VOTs, which makes the existence of a local group of neurons with bimodally distributed VOT tuning unlikely. Neurons recorded simultaneously in three different cortical areas also had nearly identical gap detection properties which does not directly support a topographic map of VOTs. Of course, a distributed non-topographic map cannot be ruled out. In addition, it is possible that categorical perception arises in neurons with both sensory and motor properties, not unlike those in HVc in birds (Doupe and Kuhl, 1999).

Although there have not been studies in auditory cortex using word-medial or word-final stop consonants, studies on detection of gaps in noise (Eggermont, 1995b, 1999) may be of relevance. Gaps with durations of 0–70 ms in 5 ms steps were placed either early in the noise, comparable to the position in a VOT discrimination task, or in the middle of the 1 s noise burst. Whereas in the early position, the minimum represented gap was typically about 40 ms in the late position, the minimum represented gap was always 5 ms. This is comparable to the psychophysical detection of gaps placed in noise (Phillips, 1999). It is also comparable to gaps in noise positioned between a leading noise burst of 800 ms and a trailing noise burst of 100 ms as measured in the forebrain (Buchfellner et al., 1989) and in AN fiber recordings of the awake starling (Klump and Gleich, 1991). Twenty percent of neurons, mostly those in field L and of the phasic-tonic type, showed thresholds between 0.4 and 3.2 ms which was comparable to the median behavioral threshold of 1.8 ms. The distribution of minimum detectable gaps in the AN fibers was similar to that in the forebrain. This suggests that the CCI and the combined interval from closure onset to release and from release to the onset of voicing in the vowel that follows the consonant would be resolved with equal accuracy. Therefore, I am not convinced that the variability that exists in the auditory periphery is of much importance in representing silent gaps in words.

8.3. Representation of AM

In humans, AM tones or AM noise produce various hearing sensations depending on the modulation frequency. These include rhythm and fluctuation strength

for AM frequencies below about 20 Hz, and roughness and pitch for AM frequencies above 20 Hz. The sensation of roughness disappears above 300 Hz and is strongest at 70 Hz, whereas periodicity pitch loses much of its perceptual strength above 3 kHz (Zwicker and Fastl, 1990). The click repetition rates and AM frequencies used generally overlap the boundary between rhythm and roughness. It is a premise that the limiting rate of phase-locked cortical responses may be related to that perceptual boundary.

The sounds that animals normally encounter in nature have broad band spectra, are highly structured in the time domain, and consist of individual segments with a duration rarely exceeding a few seconds. Temporal representation in the auditory nervous system manifests itself in two forms; the first is based on the locking of neuronal firings to the period of pure tones or to the fine structure of complex sounds, this is commonly referred to as phase locking. The second form is the locking to the, slower, AMs in a complex stimulus; a phenomenon that I will refer to as envelope locking.

The degree of envelope locking can be expressed in the tMTF. In psychoacoustic studies, the tMTF is generally expressed as a detection threshold for modulation depth, whereas in physiological studies it is given either as number of synchronized spikes or as the modulation gain as a function of modulation frequency. The tMTF is a measure of temporal acuity of the auditory system up to the point of measurement. In psychoacoustic studies it specifies the minimum detectable level of modulation, typically of a wide band noise carrier, as a function of modulation frequency. Under a variety of conditions, and for most species that are studied, the tMTF for continuous carriers resembles a low-pass filter function. This is taken as evidence that the auditory system integrates incoming information over time such that it can resolve individual amplitude fluctuations of the waveform at low modulation frequencies. In psychoacoustics there is a history to describe the auditory system with a model that involves a band-pass filter, followed by a non-linear element such as a half-wave rectifier, a low-pass filter and some central decision mechanism such as an ideal observer. tMTFs for noise carriers in normal hearing subjects (Fig. 4) are very similar to those of diverse mammalian species (macaque: Moody, 1994; chinchilla: Salvi et al., 1982). Acoustic tMTFs have maximum sensitivities at 20–80 Hz with cut-off frequencies of 55–160 Hz. The tMTF has the general shape of the low-pass filter component of the model, although the frequency slope of 3–4 dB/octave of the tMTF is less than expected from a simple first-order filter (Moody, 1994). Its similarity with $1/f$ noise, i.e. noise with spectral energy that changes with -3 dB/octave, suggests that there is an integration over bandwidth. For pure tone carrier tMTF in humans has

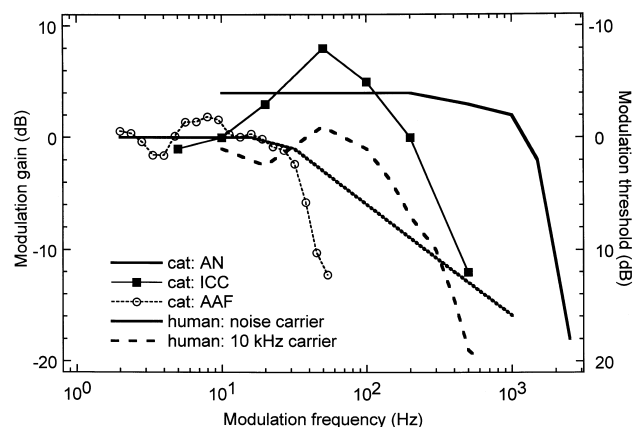


Fig. 4. tMTFs obtained electrophysiologically in cat and psychoacoustically in humans (based on Viemeister (1979) and Fassel and Kohlrausch (1995)). The human curves have been arbitrarily shifted to make their low-frequency portion equal to 0 dB gain. Three representative neural curves obtained in anesthetized animals are shown for AN fibers, for the IC and for the auditory cortex. One observes that the ICC is most predictive for the human data.

a high-frequency slope of -6 dB/octave. In AN fibers the slope is about -18 dB/octave, for VCN stellate cells it is about -12 dB/octave, and for units in the IC it approaches -6 dB/octave.

The capacity of AN fibers to follow click stimuli with various repetition rates (Kiang et al., 1965), AM tones or AM noise with different modulation rates (Frisina et al., 1990; Javel, 1980; Joris and Yin, 1992; Palmer, 1982) or random noise modulated tones (Møller, 1976a), can be described by low-pass functions with a limiting rate (the -6 dB point) of 1–3 kHz. Fig. 5 shows a compilation of tMTFs from AN to AC. Maximal phase locking is graded with SR: high in low-SR fibers, low in high-SR fibers. Maximal phase locking is also correlated with CF: it is especially low in high-SR, low CF fibers. At high CFs, a temporal limitation seems to restrict the highest f_m s to which phase locking is observed. Effects of modulation on average firing rate are small, and can be enhancement as well as suppression (Joris and Yin, 1992). It appears that tMTFs are very much like those for AN up to the lateral superior olive (LSO) in the localization path (Joris and Yin, 1992), but may increasingly become tuned to lower BMFs for the identification path. The observed modulation responses could largely be explained by a model for fast adaptation (time constant < 20 ms; Eggermont, 1975, 1985; Smith and Brachman, 1982; Westerman and Smith, 1984; Yates, 1987).

8.3.1. Localization path

Fig. 5A shows a compilation of tMTFs for the major stations in the sound localization path. All LSO afferents (spherical and GBCs, and MNTB principal cells) showed low-pass tMTFs very similar to those of AN

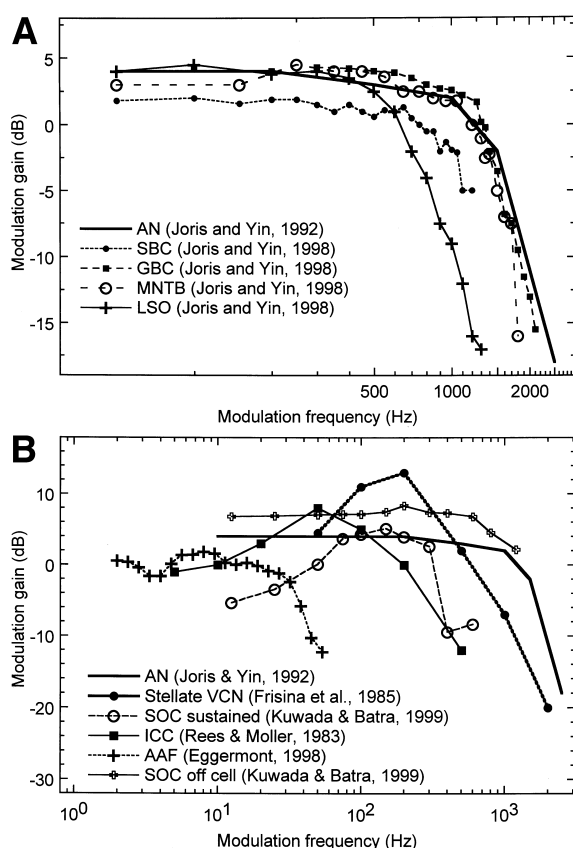


Fig. 5. tMTFs obtained in various nuclei of the localization pathway (A) and the identification pathway (B). Limitations in the capacity to follow high FM start to occur for the LSO units in the localization path. In the identification path the gradual loss of sensitivity to high FM when ascending the identification path is noted.

fibers, with cut-off frequencies that largely overlapped with the distribution for AN fibers (Joris and Yin, 1998). As we have seen, many cell types in CN show higher gains than bushy cells, but over more restricted ranges of MFs. For example CN choppers (Frisina et al., 1990) have high gains but their cut-off frequencies are well below AN values. In contrast GBCs and SBCs preserve envelope information over a wide range of MFs and this range is little affected by SPL. Synchronization of LSO cells to the envelope difference frequency in a binaural AM beat stimulus of increasing MF is restricted to MFs < 800 Hz (Joris, 1996). Because the phase locking ranges of LSO afferents are essentially AN fiber-like, the limited range for ITD sensitivity in LSO therefore must derive from post-synaptic limitations such as poor synchronization across converging MNTB or SBC afferents, from dendritic filtering or from temporal summation of subthreshold events (Joris and Yin, 1998).

8.3.2. Identification path

Fig. 5B shows a compilation of tMTFs for the major nuclei in the sound identification path. For the CN in

the rat (Møller, 1973, 1976a,b), gerbil (Frisina et al., 1990) and cat (van Gisbergen et al., 1975; Kim et al., 1990) two classes of neurons have been described with respect to the shape of the tMTF. One class showed a low-pass behavior and the other class showed a band-pass behavior. Both had BMFs in the range of 50–500 Hz. Classes of AM response could be identified with specific types classified by other means (e.g. various chopper types). In both cases the limiting rate was lower than for AN fibers and usually around 500 Hz (range 200–1200 Hz). Limiting rates and class type appeared to be independent of CF. At sufficiently low stimulus levels the change in discharge rate of the fibers and cells was linearly related to the modulation depth of the tones. Units that were band-pass at high intensity levels showed low-pass behavior at low stimulus levels (Frisina et al., 1990; Kim et al., 1990). The band-pass neurons showed a gain increase at the BMF on the order of 12 dB (Fig. 5B). Kuwada and Batra (1999) identified a class of off-neurons in the SOC, outside MSO and LSO, which showed quite different tMTFs. The sustained neurons were typically tuned in a way comparable to that of the stellate cells in VCN, but with slightly lower gain. In contrast, the off-cells showed a tMTF comparable to that of AN fibers but with higher gain. In the IC of the rat, the same two classes of tMTFs were found as in VCN. Units that were low-pass at intensities within 15 dB from threshold showed a band-pass characteristic at higher levels. Band-pass units became low-pass units in the presence of broad band noise. Both the low- and band-pass characteristics showed a steep decline in response above 200 Hz. The peak sensitivity of band-pass MTFs was invariably between 100 and 200 Hz. The period histograms had the form of a filtered and half-wave rectified version of the modulating waveform. The rectification was attributed to the low levels of spontaneous activity in the IC (Møller and Rees, 1986; Rees and Møller, 1983, 1987).

Langner and Schreiner (1988) found in the IC of the cat a very large range of BMFs for band-pass neurons; BMFs ranged from 10 to 1000 Hz although 74% of the neurons had their BMF below 100 Hz. These authors found a relative independence of CF and BMF. Most units showed a band-pass characteristic for firing rate (75%), however, the vector strength (VS) showed a band-pass character in only 33% of the neurons tested. On this basis it was proposed that in the IC a rate coding for periodicity is favored over a synchronization coding. Similar conclusions were reached for the auditory midbrain of the grass frog (Epping and Eggermont, 1986) and leopard frog (Eggermont, 1990). Snyder et al. (2000) used the number of synchronized spikes per modulation period and found basically the same results using SAM pulse trains applied through a cochlear implant in deafened cats. Müller-Preuss et al. (1994) found

that encoding of AM sounds occurred to a greater extent via phase locking of discharges than via changes in firing rate. Modulation depth appears also to be processed via phase-locked firings because spike rate on average remains constant between 100% and 0% modulation. Schreiner and Langner (1988) claimed that BMFs were topographically coded in the central nucleus of the IC. In contrast, Müller-Preuss et al. (1994) found no clear relationship between BMF and recording place or CF. Recently, in an extensive study in the IC of gerbils, Krishna and Semple (2000) found that the range of BMFs was below 140 Hz. They also found that there was no correlation of AM response properties with the BF of the neurons, and that the dependence of the rMTF on the SPL suggested that any topographic map of BMFs is likely dependent on SPL. They also noted that most IC neurons responded with significant synchrony to modulation frequencies below 300 Hz. Thus considerable temporal information remains present in the response in the range of modulation frequencies where most BMFs lie. Krishna and Semple (2000) suggested on that basis that the emergence of rate tuning does not necessarily preclude the possibility that information about modulation frequency is also present in a temporal code.

In the MGB about 70% of the units were responsive to 0.5 s long click trains, the majority thereof were normal ‘lockers’ which responded for click rates up to 100 Hz (Rouiller and de Ribaupierre, 1982). Preuss and Müller-Preuss (1990) found that the majority of neurons responded best in the range from 4 to 64 Hz, with a most frequent BMF at 32 Hz and a median BMF at 16 Hz. The median BMF was primarily independent of the quality of the signal carrier (noise or tone). With an increase in modulation depth, a majority of neurons showed a monotonic increase in the discharge periodicity whereas spike rate itself was non-monotonic in more than half of the cells. It was suggested that the total spike rate may reflect more accurately the physical energy of the stimulus, whereas the discharge periodicity may be more suited to signal the stimulus shape.

Periodic amplitude-modulated sounds such as sine or square wave modulated tone bursts could be represented differently in separate cortical fields in cat (Schreiner and Urbas, 1988). The average BMF in AAF of the paralyzed and lightly anesthetized cat was higher than in AI. This was largely the result of higher BMFs (up to 100 Hz) for units with CFs above 10 kHz, whereas for lower CFs all BMFs were below 20 Hz and similar to those in AI. In AII a range of ‘normal’ BMFs close to 10 Hz was found but supplemented by a large group of BMFs with values below 5 Hz. Such low BMFs were less frequently found in AI or AAF (Schreiner and Urbas, 1986). In agreement with the

findings of Schreiner and Urbas (1988), AM tones showed poor tMTFs in AII. However, simultaneous recordings from AI, AAF and AII in cats showed no significant differences between tMTFs obtained for periodic click trains and AM noise (Eggermont, 1998c). In the awake squirrel monkey, 78.1% of all acoustically driven neurons encoded the envelope of the AM sound; the remaining 21.9% displayed simple on, on/off, or off responses (Bieser and Müller-Preuss, 1996). 46.6% of all neurons had the same filter characteristic for both the overall firing rate and the envelope-locked response. Varying modulation depth was encoded by the neuron’s ability to follow the envelope cycles and not by the spike rate frequency. AI, the first temporal field (T1), and the para-insular auditory field (Pi), which showed phase locking in the range of 2–64 Hz, showed the best temporal resolution. The anterior field in particular failed to follow the AM envelopes. In the insula, thought to be involved in the exercise of language and speech perception in humans, the high proportion of neurons unable to encode AM sound (40.6%) and the low mean BMF (9.9 Hz) did not support a prominent role in temporal information processing. Many of the auditory responsive units in insula showed remarkably short latencies, an indication that the main input to this region is not from other cortical regions. This is in keeping with the findings that the supra-geniculate nucleus projects to the insula. The neural data in the most responsive fields and the AM frequency range of squirrel monkey calls (4–64 Hz) suggest a correlation between vocalization and perception. In the forebrain of the awake mynah bird (Hose et al., 1987) 65% of neurons in input layer L2 of field L preferred AM rate above 20 Hz (median 35 Hz; 2 S.D. range 5–254 Hz). In contrast, in the layers above and below the input layer 85% of units showed preference for periodicities below 20 Hz (median 7.8 Hz; 2 S.D. range 1–55 Hz).

Comparison of the effect of various periodic stimuli such as sinusoidal AM and rectangular AM tones (Schreiner and Urbas, 1986, 1988) with periodic click trains (Eggermont, 1991, 1993a; Schreiner and Raggio, 1996) in cat AI suggests that the BMFs are highest for sinusoidal modulation and lowest for periodic click trains. Phillips et al. (1989) used continuous repetitive tone pips and analyzed them in terms of the response per tone pip, i.e. in terms of entrainment. Eggermont (1991) showed a similar analysis for periodic click stimulation and Schreiner and Raggio (1996) also presented their data in this alternative format in addition to using the tMTFs. These entrainment functions are low-pass functions of click (tone pip) repetition rate and were very similar in these three studies (Eggermont, 1997a,b). In rat AI, the average repetition rate transfer function for 25 ms tone bursts was low-pass with a sharp drop-off in evoked spikes per tone above 9 Hz

(Kilgaard and Merzenich, 1999). Lu and Wang (2000) using interclick intervals as short as 3 ms found that the median synchronization boundary in the cat was at about 25 Hz. For higher repetition rates, but not above approximately 160 Hz, firing rate was assumed to represent the click-repetition rate. They suggested that for repetition rates below approximately 100 Hz a temporal representation was likely and for repetition rates above 100 Hz more information was available in the firing rate. An alternative way to arrive at tMTFs is the use of moving ripple stimuli (Kowalski et al., 1996). They found that the temporal and ripple transfer functions were separable. This implies that the combined spectro-temporal transfer function of a cell can be written as the product of a purely ripple transfer function, exhibiting the spatially mapped response properties, and a purely temporal transfer function. The ripple parameters that characterized cortical cells were distributed somewhat evenly, with the characteristic ripple frequencies ranging from 0.2 to >2 cycles/octave and the characteristic angular frequency typically, corresponding to BMFs, ranging from 2 to 20 Hz.

All results obtained for AM stimuli in auditory cortex using period histograms point to significant synchronization up to 50 Hz (Fig. 5B), with a potential temporal code up to frequencies of about 20–30 Hz (Eggermont, 1991; Gaese and Ostwald, 1995; Schulze and Langner, 1997). Clear evidence for such an internal temporal code based upon interspike intervals has been demonstrated in cortex for the AM frequency range below 32 Hz (Eggermont, 1998c), recently upgraded to 64 Hz (Eggermont, unpublished observations). Psychophysical data (Fig. 4), expressed as a modulation threshold and inverted to fit the modulation gain scale, suggest a better performance for pure tone carriers than for wide band noise carriers.

8.4. Representation of FM

In general the sensitivity to sinusoidal AM and FM rate is the same in AN fibers (Sinex and Geisler, 1981) and in CN neurons (Møller, 1972; Vater, 1982), and such neurons also respond to pure tones. The preferred velocity for linear FM sweeps in CN was around 30–100 kHz/ms (Møller, 1974). The IC is the first station in the auditory pathway to feature neurons that respond only to FM and not to pure tones (Poon et al., 1992). Such FM-specialized cells (about 34% of total) had larger dendritic fields, more dendritic branching and more dendritic spines than did cells which also responded to pure tones (43% of total). Best FM velocity for mixed cells was about 16 octaves/s and about 33 octaves/s for FM-specialized neurons. Preferred FM range was about 0.7–1 octaves for mixed cells and one to two octaves for FM-specialized cells. With click

stimuli, FM-specialized cells were activated at a longer latency (about 2 ms, 10 vs. 12 ms) than mixed cells.

A number of investigators have provided evidence suggesting that FM and AM stimuli are processed independently by parallel pathways or channels in the auditory system (Kay, 1982; Regan and Tansley, 1979). Gardner and Wilson (1979) show evidence for direction of FM-specific channels in addition to a repetition rate-specific channel. However, Wakefield and Viemeister (1984) have questioned the interpretation of these adaptation studies and suggest that the results could also point to effects of non-sensory factors such as the use of different strategies. Saberi and Hafter (1995) note that there is an FM to AM transduction when the FM sweep passes through the frequency tuning curve. They found that one can perceive a single image in the head depending on the phase difference between an FM sound (e.g. 250 Hz modulation of a 3 kHz carrier) presented to one ear, and an AM sound, with the same modulation and carrier, to the other ear. This suggested that FM and AM sound must be transformed into a common neural code in the brainstem, at least as far as the localization pathway is concerned. Gaese and Ostwald (1995) showed that in auditory cortex of rats the BMFs, for sinusoidal AM and FM, were significantly correlated for the two stimuli. This suggests that there is also a common mechanism for repetition rate, and that there is a pool of units that is activated by both AM and FM stimuli. This was also found in cat AI (Eggermont, 1994a).

FM studies are generally thought of as relevant in understanding formant transitions in speech. However, as Heil and Irvine (1998a) have noticed that such sounds, although important, constitute only a very small proportion of the acoustic biotope (Aertsen et al., 1979) that animals can discriminate and identify. The correspondence between the apparent tuning properties of neurons in species such as rats and cats and the equivalent acoustic parameters in such vocalizations is generally poor. For instance, a kitten's isolation call is characterized by a rise and fall of at least five harmonically related frequency components over a 1 s period (cf. Fig. 1). The maximum rates of change of frequency (RCFs) are below 0.032 kHz/ms for the fundamental and first three harmonics. The FM components of other calls fall within similarly low-frequency and RCF ranges. Only a very small proportion of the multi-units (MUs) sampled had best RCFs matching this low RCF range of the cat's vocalizations, however, was sufficient to cover the range of RCFs in natural vocalizations that may be behaviorally relevant for the cat. Analysis of a database of natural sounds showed that this is also true for vocalizations of other animals, e.g. songbirds (where most of the FM components are slower than 0.08 kHz/ms). It is likely that the large proportion of neurons

sensitive to fast rates of FM is also activated by other transient broad band signals, such as would be produced by prey and objects moving in natural environments (Heil et al., 1992a). In contrast, velocity preference in the ferret auditory cortex seems to be better correlated with the natural range of FM velocities. This finding may imply that the auditory cortex of ferrets is better tuned to FM sounds occurring in the environment than the cat auditory cortex (Nelken and Versnel, 2000).

Using linear FM ramps of 2 kHz excursion and varying CF, Phillips et al. (1985) found that directional preference was strictly associated with the low- and high-frequency slopes of the cell's spike count vs. frequency function (iso-intensity contour) over the frequency range covered by the ramp.

In cat AI, the response functions to unidirectional FM sweeps of varying linear RCF that traversed the excitatory frequency response areas (FRAs) displayed a variety of shapes (Heil et al., 1992a). Preferences for fast RCFs (> 1 kHz/ms) were most common. Best RCF was not correlated with sharpness of frequency tuning. About 2/3 of the MU responses showed a preference for downward sweeps. This is in contrast to the findings in rats, where units responded to sinusoidal FM stimulation in 52% for the upward direction, in 30% for the downward direction, and 18% had no directional preference (Gaese and Ostwald, 1995). This may be a species effect because Eggermont (1994a) found that all but a few units in cat AI responded only to downward modulation sweeps. DS was independent of CF and independent of best RCF. Discharges of a given MU over its effective RCF range were initiated at the same instantaneous frequency (effective F_i) independent of RCF. Effective F_i values fell within the excitatory FRA of a given MU. Responses were evoked only when the frequency of the signal was modulated towards CF (either from above or from below) and not when modulated away from it. Responses were initiated before the modulation reached CF. Stimulus intensity had only a moderate effect on RCF response characteristics and DS.

In another study (Mendelson et al., 1993) exponential (i.e. log frequency varies linear in time) FM sweeps from 0.25 to 64 kHz were presented at three rates of change (110 octaves/s, 33.3 octaves/s, 10 octaves/s). The high-frequency domain of AI (3.2–26.3 kHz) was mapped over its dorso–ventral extent. Neuronal responses (506 single-unit and MU responses in six animals) for frequency sweeps of different speeds appeared to be systematically distributed along the dorso–ventral axis of AI. In the dorsal region, cortical cells responded best to fast and medium FM sweeps, followed more ventrally by cells that responded best to medium then slow then medium-speed FM sweeps. In the more ven-

tral aspect of AI (may have contained cells from AII), neurons generally preferred fast sweeps.

DS appeared to be non-randomly distributed along the dorso–ventral axis of AI. In general, units that responded best to upward directed FM sweeps were located in the more dorsal and ventral aspects of AI, while units that responded best to downward-directed FM sweeps were usually located in the mid region of AI. Units that responded best to slower FM speeds also seemed to have sharper tuning and vice versa. Responses to broad band transients were also systematically distributed across the cortex such that those units most responsive to clicks were located in the more dorsal and ventral subregions. It was found that the distribution of preferred speed and direction responses for FM sweeps were correlated with integrated excitatory bandwidth and broad band transient responses. Neither the preferred speed nor the preferred direction appeared to change significantly as a function of cortical depth, especially for responses encountered more than 650 μ m from the cortical surface. In general, the largest group of MU responses displaying a unimodal rate tuning profile responded best to fast (110 octaves/s) FM sweeps, the second largest group preferred medium (33.3 octaves/s) FM sweeps, and the smallest group preferred slow (10 octaves/s) sweeps. For linear FM sweeps, Heil et al. (1992b) found that best RCF was distributed non-randomly along isofrequency strips, although there was considerable variability between cats. In contrast, FM DS and sensitivity were randomly distributed.

Because of the apparent different findings in linear versus exponential FM sweeps, Nelken and Versnel (2000) compared exponential and linear FM speeds in the same cats and found that the neuronal preference for FM velocity was independent of frequency trajectory. Thus clusters that responded best to slow logarithmic FM also preferred slow linear FM and vice versa. Consequently, topographic distributions of velocity preference were roughly independent of the stimulus paradigm. However, topographic distributions of DS differed between linear and logarithmic paradigms. The most frequent preferred velocity for log sweeps was 30 octaves/s, followed by a low percentage flat distribution up to 300 octaves/s. For linear sweeps the preferred rate was 0.2 kHz/ms, with the distribution slowly decreasing to 2 kHz/ms.

The differences between the cat and ferret may point to real species differences. The first is the velocity sensitivity for linear FM sweeps. Heil et al. (1992a) reported an almost exponential increase in the number of clusters preferring fast FM velocities in cat AI. Tian and Rauschecker (1994) show a preponderance of neurons that prefer fast sweeps in cat AAF. In contrast, there is a much larger representation of slow FM

velocities in the ferret. In particular, the histogram of velocity preference decreases with increased velocity, the opposite pattern to the results of Heil et al. (1992a) and Tian and Rauschecker (1994) in the cat. In contrast, velocity preference for logarithmic sweeps seems to correspond better to the data in the cat. The second difference is the bias in DS. In the cat, the preferred sensitivity to downward FM sweeps is consistently reported (but only in the middle area of AI!!). In the ferret, this bias is absent, and in fact the opposite preference is common. This suggests that because of the dependence on the exact stimulation paradigm, functional maps of DS are not an invariant feature of cortical organization. Furthermore, this organization may be species-specific, and in addition there may be large differences even between animals in the same species (Nelken and Versnel, 2000). Therefore, the organization of response areas suggests a model in which the shape of the spectral envelope is encoded explicitly in the two-dimensional distribution of neural activity across the extent of AI. This could be interpreted as if the location of the maximal response along a particular isofrequency plane represents the locally smoothed gradient of the spectrum at that frequency (Shamma et al., 1993; Horikawa et al., 1998). This mapping scheme provides a neurophysiological basis for the significant role that spectral gradients play in timbre perception and the recognition of complex sounds (Plomp, 1976).

Thus, AI and AAF in cat have similar FM sensitivity with a preference for high rates of change, which is not much different from that in IC or CN. In contrast, for the posterior field (P) in the cat (Heil and Irvine, 1998a; Tian and Rauschecker, 1998), the RCF response functions were generally of low-pass characteristic or peaked at rather low RCFs (<0.1 kHz/ms). Neurons with strong responses to high RCFs all had monotonic or weakly non-monotonic rate intensity functions. RCF response functions and best RCFs often changed with SPL. Individually DS also changed with RCF and SPL, but the distribution of DS across all neurons was independent of RCF. Together with studies of responses to AM, and tone onsets (Kilgaard and Merzenich, 1999), these findings suggest more generally that field P may be best suited for processing signals that vary relatively slowly over time such as vocalizations.

8.5. *Vocalizations and other natural sounds*

In the 1970s the prevailing hypothesis was that species-specific vocalizations were represented by the activity of specific neurons: the call detectors. However, individual neurons in the auditory cortex were often found to respond to more than one call or to various features of calls (Newman and Wollberg, 1973; Winter and Funkenstein, 1973; Sovijärvi, 1975). In field L of

the European starling, Leppelsack (1978) found about 30% of neurons that responded highly selectively to individual calls, however, in a later study (Müller and Leppelsack, 1985) this percentage was considerably downgraded. The call detector neuron hypothesis was replaced by the idea of encoding complex sounds by the discharge patterns of neuronal populations (Creutzfeldt et al., 1980). This has been elaborated into neural assembly hypotheses for encoding any behaviorally relevant complex sound. Whereas the first hypothesis could be easily tested, the problem with the second is that it cannot be tested. All that can be tested is that individual neurons are responding to a large number of complex sounds and that distinguishing complex sounds on the basis of population responses works better the larger the population. It has to be explicitly tested whether population firing rate, population synchrony or the order in which a population is activated represents the complex sound.

It is also generally assumed that complex sounds are a mix of AM and FM components coupled with a harmonic complex and some wide band noise parts. Often the neurons respond to a complex sound in a combination sensitive manner, i.e. the addition of activity to parts (either in the time or frequency domain) of the sound is non-linear (Rauschecker, 1998). The complex sound can be producing more activity than the sum of its components (this is what is usually reported in the literature), but there is no reason why the response to the entire sound would not be less than the sum of the responses to individual components. In general reversal of a call has little effect except under certain conditions. Sometimes species-specific vocalizations have AM components that are in the preferred range of cortical neurons. One example is the marmoset twitter call with a natural phrase repetition rate of about 8 Hz (Wang et al., 1995). It is no surprise that this call provokes good phase locking to the envelope, and because of its nature (each phrase has a rapid attack and a slow decay) causes much larger firing in its natural order than does a time-reversed call. (This contrasts with the response of HVC cells to the bird's own song which has a slow attack and fast decay but is still preferred (Theunissen and Doupe, 1998)). However, recently Wang (2000) reported that in cat auditory cortex the preference for natural order marmoset twitter calls does not exist and that the preference found in marmosets reflects species-specific detection mechanisms. Also recently, Eggermont (2001) reported that in cat AI the neural synchrony to gamma tones is far greater than to time-reversed gamma tones, albeit that the firing rates evoked by the two tones are the same. This was specifically reflected in the tMTF, where the VS, BMF and limiting rates were significantly lower for time-reversed gamma tones. This would suggest a strong sen-

sitivity to natural and time-reversed calls of the twitter type.

Another example is the twitter call in squirrel monkeys with a 12.4 Hz repetition rate in addition to a bi-directional FM in each call syllable. Neurons in primary and rostral auditory fields as well as the insula respond strongly to the combined AM and FM of the natural calls. The insula, which increased dramatically in size during primate evolution, is a region where sensory information joins with information about the motivational state of the animal. This response in all three areas is better than that to synthetic periodic FM stimuli alone; the phase locking to these stimuli furthermore starts to decrease rapidly above 16 Hz repetition rate (Bieser, 1998). Similar phase-locked responses have been found in guinea pig cortex to song bird twitter sounds (Creutzfeldt et al., 1980) suggesting that it may have less to do with behavioral importance than was suggested by Wang (1998). This preference for the natural order of the call has also been found in field L of the starling (Leppelsack, 1978) but not in field L of the zebra finch, whereas it was pronounced in the fore-brain song nucleus, HVC (Lewicki and Arthur, 1996). Similar preferences for phase locking were reported for the twitter call in the squirrel monkey which has a slightly higher AM rate (~ 10 Hz) than that in the marmoset. In contrast, the err-call with its much higher phrase repetition rate (20–25 Hz) rarely caused envelope locking of spikes (Newman and Wollberg, 1973). Creutzfeldt et al. (1980) also observed that auditory cortical neurons responded to fewer components of a call than thalamic neurons, even if the cortical and thalamic cells were synaptically connected.

Vocalizations in the cat (Fig. 1) do have a strong repetitive component, albeit much faster than in human vocalizations and too high for envelope following. They are characterized by a harmonic complex of about six components that is slowly amplitude- and frequency-modulated. Consequently, time reversal of the call does not have much effect on the response (Sovijärvi, 1975). Cat vocalizations largely excited the cells only at the onset or offset of the stimulus sound and caused inhibition or no response at all during the other parts of the sound, the pattern thus resembling the response to pure tones (Sovijärvi, 1975). In a recent study (Gehr et al., 2000) we found a subset of cortical neurons, called peak-tracking neurons, that locked to strong features in the envelope and these vocalizations and as a result were very sensitive to time reversal of vocalizations. A sustained response and a clear tMTF to AM tones characterized these neurons which comprised about 40% of the total. In contrast, the remaining neurons only responded to the onset of the vocalization and also only to the first modulation period of the AM tones. Interestingly, the tMTF to periodic click

trains did not discriminate between peak-tracking neurons and onset responders.

9. Transformation of representations along the auditory pathway

Single-unit response properties change considerably from AN to auditory cortex. AN fibers have response functions that are monotonic, have relatively narrow frequency tuning curves, and can follow carrier frequencies up to 5 kHz and AM frequencies up to 3 kHz. Depending on the cortical area, response functions can be dominantly non-monotonic or monotonic, frequency tuning can be narrow or broad, single- or multi-peaked, phase locking to pure tones is absent and that to AM ceases below 64–128 Hz, depending on the anesthesia/alertness state of the animal. These changes are the result of various neural transformations and are hypothesized to serve at least three functions: creating level tolerance of responses, creating efficient population representations of acoustic features and to induce a shift of emphasis for representing change in the acoustic biotope.

9.1. Transformations create level tolerance

One could assume that the auditory system evolved to improve the signal to noise ratio under natural listening conditions. Several adaptations accomplish this: spatial filtering by virtue of binaural hearing, spectral filtering by virtue of narrow frequency tuning curves, synchrony suppression by dominant formants and temporal filtering resulting in enhanced AM representations. All are particularly well suited to enhance speech perception in background noise.

One of the important features of auditory perception is that it is level tolerant. Speech sounds are perceived very much the same regardless of the level or the signal to noise ratio, and categorical perception boundaries do not shift appreciably with level. Pitch is not very much level-dependent, critical bandwidth is level-independent and so is frequency discrimination. Sound localization is also level-independent. Yet, in the AN, the spectral representation of vowels in the firing rate of neurons with CFs around the formant frequencies is level-dependent. Due to the mechanical properties of the cochlea, AN fibers do not respond to a single frequency at supra-threshold levels, but rather to a range of frequencies that broadens considerably as sound intensity is increased. So at the sound level of normal conversation, virtually all AN fibers are continually active. Thus, the spike rate of a neuron on its own cannot provide unambiguous information about either frequency or intensity of sound. This ambiguity can only be resolved by

comparing relative amounts of activity in different units (Covey, 2000). Temporal coding of frequency through phase locking will of course provide a level tolerant representation but only for frequencies below 4–5 kHz.

Neural transformations that occur between AN and IC cause the neural representation of perceptual important parameters to become level-independent. This can already be found in the vowel spectrum representation through firing rate at the level of the stellate (chopper) cells in VCN. Rate intensity functions averaged from neurons with similar CF, or across all neurons recorded from, in cat ICC showed a constant average discharge rate over at least 60 dB (Ehret and Merzenich, 1988a,b). This was interpreted as excluding sound intensity coding by means of average neuronal discharge rate in the ICC. However, it could well be the first step toward level-independent coding of sound features. ICC neurons in cat also have a level-independent critical bandwidth (Ehret and Merzenich, 1988a,b). In contrast, my findings (Eggermont, 1989) in the auditory midbrain of the leopard frog suggested a maximum dynamic range of about 60–70 dB, both for the single unit and the population. Synchrony of the firings with the stimulus, and with the firings of other neurons, was largely independent of stimulus level. This suggested that in the auditory midbrain of the leopard frog sound intensity can be coded by firing rate. It may well be that sound intensity is an important parameter in anuran communication. Onset latencies, in particular in VNLL on-cells, are the same over a wide range of intensities (Covey and Casseday, 1999). The level independence of sound localization could be based on level tolerant coincidence detection in the nucleus laminaris of the owl (Pena et al., 1996).

The synchronization MTFs, comparable to those based on VS, appear to be relatively insensitive to stimulus level (between 35 and 85 dB SPL) for both click train and AM noise burst stimulation (Eggermont, 1994a). In contrast, rMTFs and tMTFs, based on total firing rate respectively synchronized firing rate, are strongly intensity-dependent.

A transformation that could result in level tolerant frequency discrimination could be based on the tendency to keep frequency tuning curve bandwidths narrow by lateral inhibition, despite the broadening effect of converging/diverging connections. However, in an intriguing study, Phillips et al. (1994) found that whereas at low SPLs (10–20 dB SPL) small foci of activity occurred along the isofrequency line representing the test frequency, at middle and high SPLs, the CFs of neurons activated by a pure tone varied across three octaves. At the highest level used (80 dB SPL), the active neurons were widely spread across the AI, and the distribution of activity had a pattern bearing little relationship to the threshold CF contour map. These

findings indicate that only isolated patches of units within the relevant isofrequency contour are activated by a given supra-threshold contralateral tone. In contrast to this single-unit study, intrinsic optical imaging of auditory cortex, which is based on membrane potential changes, shows a contiguous area of activity at high intensities (Harel et al., 2000). Thus, as Phillips et al. (1994) suggested the break-up in patches may be due to the non-monotonic nature of individual neuron's rate intensity functions. The spatial neural firing rate representation of a stimulus delivered to the contralateral ear appears, therefore, to be highly level-dependent and discontinuous. This begs the question how perceived pure tone pitch can be level tolerant, and can be discriminated from other tones in a level-independent fashion.

Pertinent to this question is whether stimulus intensity is always the relevant parameter for discussing level tolerance. Whereas for steady-state sound, the intensity level is the correct parameter, for transients that make up a large part of vocalizations, it has been argued that the maximum acceleration of peak pressure is the appropriate parameter (Heil, 1997a,b; Heil and Irvine, 1997). It was found that first-spike latency was closely related to this maximum acceleration under a large range of rise times (for a cosine squared envelope) and peak SPL values. This effect was already present in AN fibers but is still robust in auditory cortex. The implication of this is that the integration window of the neuron, starting at stimulus onset and ending close to first-spike latency, is variable and level-dependent. This in itself could produce a level tolerant effect. Combined with adaptation and suppression effects for the steady-state aspects of vocalizations, which are also stronger for higher intensities, a number of mechanisms may contribute to achieving level tolerance.

Potentially, a transformation that makes responses somewhat resistant to background noise has been inferred from activity in the central region of AI, where the effect of background noise is limited to a shift in the threshold and rate intensity function (Ehret and Schreiner, 2000). Natural sound backgrounds show strong co-modulation which makes it less effective in masking a foreground sound (Nelken and Versnel, 2000), so co-modulation masking release is another mechanism that creates tolerance to background noise masking.

9.2. *Transformations enhance tuning and so increase the efficiency of population coding*

Another type of neural transformation is potentially found in a sharpening of the tuning to a biologically important parameter of sound. An example is formed by the tuning for interaural time differences. The tuning

curve of neurons in the MSO peaks for a given ITD, reflecting the location of the sound in azimuth. However, the half-width of the tuning curve is appreciable; in rabbit it is on average 730 μ s. Thus, accurate spatial tuning will require a population mechanism. However, the ITD tuning curves are much sharper at the level of the ICC with a half-width of 430 μ s, and an additional sharpening occurs in the thalamus to reach 350 μ s (Fitzpatrick et al., 1997). This, presumably final, sharpening still is more than half of the sensory ITD field for the rabbit which based on head size would be about 600 μ s. So population coding remains essential. Fitzpatrick et al. (1997) calculate on the basis of a population vector model that the number of neurons required to reach a certain acuity decreases sharply with decreasing half-width of the ITD tuning curve. For example, using 40 neurons a change of 147 μ s could be detected in MSO, a change of 39 μ s in the IC, and 16 μ s in the thalamus. Further sharpening, however, does not necessarily result in a further increase in efficiency as artificial sharpening by another 50% has shown. This suggests that the tuning of thalamic neurons is nearly optimal for maximum efficiency for the accurate representation of ITDs using a minimum number of neurons. The auditory system therefore sharpens overly broad tuning from the MSO until it is appropriate for a maximally efficient population code.

9.3. *Transformations emphasize change in stimulus content*

Another transformation is that from sustained responses into responses that reflect changes in the stimulus content on relatively slow time scales (Covey and Casseday, 1999). This results in an increase in the percentage of on, off and on/off responses along the auditory pathway up to the cortex. Such transformations also result in the loss of phase locking to the carrier frequency and in the gradual loss of AM following for modulation rates above 100 Hz. The subthreshold activity of each individual target neuron that receives input from the population, can be seen as changes in the membrane potential or synaptic currents that are summed into LFPs, reflects the net effect of input from the entire population. The target neuron will produce an output only if the timing and nature of the combined input from the population are such that the target neuron is depolarized strongly enough and quickly enough to reach threshold. This requires synchronized input from a relatively large number of input neurons.

Strictly speaking, discrimination of two successive sound parameters reflects the difference between two representations evoked in the same or a different group of neurons. Perception of stimulus differences in both

humans and animals is strongly correlated with the amplitude of various surface-recorded evoked potentials (EPs). An EP in response to a stimulus represents the difference in neural activity evoked by that stimulus and the neural representation of another stimulus (or the absence of a stimulus for that matter) that is kept in memory. So, EPs signal perceptual or behaviorally relevant differences in neural activity evoked by the current stimulus and a previous stimulus or by stimulus change. Behavior generally results as a consequence of similar differences.

Because the scalp recorded EPs have an established relationship with sensation, perception and cognition, it would be nice if we knew the relationship between these EPs and the activity at the level of single nerve cells. For the auditory brainstem response this is not too difficult as it represents the compound action potentials from successive generation sites in the AN and brainstem (Eggermont and Schmidt, 1990). So one expects good correlations between the lowest thresholds for neurons in each of these stations and the threshold of hearing (Liberman, 1978).

Most electrophysiological recordings from auditory cortex find short latency LFPs accompanied by short latency spike activity (Eggermont and Smith, 1995). Except for rebound activity (Eggermont, 1992b), generally no neural correlates of long latency components such as N1 (100 ms latency) are present, at least not in the anesthetized animal and only in a few long latency neurons in the awake animal. So, the interface from single cell activity to behavior (perception) via the macro potentials recordable from the scalp may provide new avenues for assessing the neural code. EPs can only be detected if a sufficient number of neurons is activated synchronously, with sufficient speed to provide for a measurable derivative, and if the electrical fields of these neurons do not cancel each other. Macro potentials, by their very nature of being a compound temporal derivative of the membrane potential of many cells, only can signal change in the stimulus. They are thus mostly used in discrimination tasks and may represent the various stages leading to conscious decisions.

Recording neural responses to speech sounds in AI of awake monkeys (Steinschneider et al., 1994) using concurrent multi-laminar recordings of auditory EPs, the derived current source density, and MUA suggested some interesting transformation of the response to syllables /da/ and /ta/ between thalamus and cortex. They found that 77% of thalamo-cortical fiber responses were phase-locked to the 100 Hz periodicity in the syllable, 15% showed a 'double-on' response bracketing the VOT, and 8% showed another pattern. Cortical MUA (lower layer III) showed that only 45% were phase-locked, while 44% showed a 'double-on' re-

sponse, and 11% were other type responses. The double-on pattern was found most often in low CF regions of the cortex. In contrast, the phase-locked pattern occurred most commonly in the high CF region of AI. Cortical ‘double-on’ responses were significantly more common than in TC fibers, while in the TC fibers, the phase-locked pattern predominated. Thus AI transforms the phase-locked temporal pattern into a pattern that accentuates the onset of the aperiodic and periodic segments of the syllables. This loss of temporal dynamics has as a direct consequence that temporal synchrony among many active neurons is increased, especially at points of change in the stimulus.

10. Is texture mapped, do contours synchronize?

Some of the auditory features, for which we reviewed the neural representation at various stations along the auditory pathway, include common onset and offset of sound, common rates of AM and FM, harmonicity and common spatial origin. It is useful for the subsequent synthesis to reemphasize that these auditory features can be categorized into contour and texture components of sound. Contour components are those temporal aspects of sound that covary across frequency, overlap with IBPs, and are likely exclusively coded in the temporal domain. Onsets, noise bursts and common rates of slow (< 20 Hz) AM and FM, i.e. the region where rhythm is dominant, are clearly contours that delineate for instance sound duration and separation between noise bursts and formant transitions. Common higher rates of AM, in the roughness and pitch range, common changes in formant frequencies, and harmonicity, i.e. IBEs, are aspects of sound texture. Texture aspects can be further characterized by pitch, timbre and roughness and slow changes therein. These texture aspects of sound thus relate to, constant or slowly changing, spectral representations in a cortical rate-place code. Pitch, timbre and roughness are thus reflections of the sound representation across different cortical areas. One expects texture and contour components thus to have largely independent neural representations in auditory cortex or potentially also in subcortical areas.

The temporal representation of rhythm as well as sound onsets is strong in both primary cortical areas, AI and AAF, whereas it is distinctly less punctuate or absent (for AM tones) in AII (Eggermont, 1998c). Thus the primary areas may be more suited to represent all the contour aspects of sound. Specifically, because of its strong response to clicks and noise onsets, AI may be specialized in sound transients such as onsets and offsets (Phillips, 1995). In AII, low-frequency AM of broad band signals, clicks and AM noise are still effec-

tively represented in temporal code, but slow changes in the AM of pure tones are not well represented (Eggermont, 1998c; Schreiner and Urbas, 1988).

One of the strongest synchronizing features of sound is stimulus onsets, they show synchronization coefficients in response to clicks and noise bursts of about 0.125–0.15 for single electrode pairs, and about 0.05 for dual electrode pairs in AI (Eggermont, 1994b), as well as between auditory cortical areas. The post-stimulus control values for single electrode pairs within auditory fields are about 0.08, and for dual electrode pairs in AI about 0.05, and for pairs in different cortical fields about 0.015 on average (Eggermont, 2000a). Thus the relative increase of the synchrony produced by stimulus onsets is larger between cortical areas than it is within a cortical area.

In the data reviewed here, evidence was found to suggest that the texture of sound is represented in topographic maps in AI, mainly along the isofrequency sheets but in combination with the tonotopic map. Texture is coarsely mapped, but a very sensitive local representation could coexist in these maps. This could be based on a ratio code not unlike that for color in the visual system (Erickson, 1974). A similar code could exist in every topographic map but requires a local representation of at least three small groups of neurons differently tuned to the feature mapped, e.g. to periodicity pitch. As a consequence, the texture map boundaries cannot be smooth. In case of periodicity pitch, the demonstrated very fine pitch resolution would then be formed in different local populations by the relative firing rates of neurons broadly, but differently, tuned to the different pitches.

Because the large majority of units in AI, AAF and AII also respond in temporal and synchronous fashion to the dominant contours (Eggermont, 2000a), this synchrony can bind the place-represented texture of a complex sound (Ohl and Scheich, 1997). Recent experimental data (Stevens and Zador, 1998) suggest that only large inputs, indicative of coordinated bursts from many input cells, can generate the large current fluctuations requisite for the experimentally observed spike output irregularity in cortex. As a consequence, temporal coordination of inputs may be an important determinant in cortical signaling mechanisms and could be important for representing the steady-state portions of sounds as well. This was shown recently for both pure tones and steady-state noise (deCharms and Merzenich, 1996; Eggermont, 1997a,b). As further evidence for this numerous synchronized input scenario, a strong link between the spontaneous coordinated input activity of many cells (the LFP) and the single-unit activity recorded on the same electrode, which was similar to that under stimulus conditions, was previously found in auditory cortex (Eggermont and Smith, 1995).

11. Proposal for a neural code: synchronized topographic maps

We have reviewed evidence for a place representation of texture features of sound, such as frequency, periodicity pitch, harmonicity in vowels, direction and speed of FM, and for a temporal and synchrony representation of sound contours, such as onsets, VOTs, and low rate (< 32 Hz) AM. For most of these proposed representations, the evidence is not overwhelming or only provided by one laboratory. For instance, for the topographic maps for periodicity pitch and vowel primary frequency differences there are only one or two published accounts each. In contrast, the changing properties of neurons along the medial–dorsal extent of AI are well documented and related to preferences for FM direction and speed, integrative properties across frequency, and latencies. The temporal representation of stimulus onsets and of low-frequency AM is well documented, as is the onset and steady-state response synchrony across recording sites within AI. In contrast, the across cortical area synchrony has only been provided in one, very recent, report from this author's laboratory. One also needs to keep in mind that studies in awake animals (largely birds and monkeys) suggest that periodicity up to a few hundred Hz can be coded in temporal fashion. This high upper value is likely limited to the cortical input layers and most studies do suggest that the upper limit in the upper and lower cortical layers is at most 100 Hz. This does not mean that the temporal coding in cortical input layers is actually used, e.g. for coding voice pitch, but it cannot be excluded.

Given these caveats, it is logical to assume a multiplexed coding of complex sound whereby the contours set up widespread synchrony across neurons activated by the texture of sound in all auditory cortical areas. Presumably, albeit with only scant evidence for cat and monkey, these different cortical areas extract and map different sound texture features. Contours mark changes and transitions in sound and auditory cortex appears particularly sensitive to these dynamic aspects of sound. Texture determines which neurons, both cortical and subcortical, are activated by the sound whereas the contours modulate the activity of those neurons. Such a synchronized map code would be fairly resistant against the change or break-up of topographically maps with changing intensity because of strong redundancy in those maps, and also because only the synchronized parts matter. Such a code would also be resistant against filling in short stretches of the texture by noise because most of the mapped areas will respond to noise as well as they do to harmonic complexes or pure tones. This emphasizes the, above reviewed, primary impor-

tance of temporal (contour) aspects of sound over the more static, texture aspects in perception.

Because contours are temporally represented in the majority of neurons activated by the texture aspects of sound, each of these neurons is part of an ensemble formed by the combination of contour and texture sensitivity. Because nearly all neurons will respond to contours but are differentially tuned to the (mapped) texture, the latter determines the potential extent of the ensemble. However, the most important parts of sound for perception are the contours as is dramatically illustrated by the sound features that are preserved for stimulation in cochlear implants. Thus, the temporal and synchrony coding that represents those contours is likely to be the dominant part of a neural code. It is exactly this part that is also reflected in the ERPs that appear so well matched with the quality and discrimination of sounds in sensation and perception.

Acknowledgements

This investigation was supported by grants from the Alberta Heritage Foundation for Medical Research and the Natural Sciences and Engineering Research Council of Canada, and by the Campbell McLaurin Chair for Hearing Deficiencies.

References

- Abeles, M., 1982. *Local Cortical Circuits: an Electrophysiological Study*. Springer Verlag, Berlin.
- Abeles, M., 1988. Neural codes for higher brain functions. In: Markowitsch, H.J. (Ed.), *Information Processing by the Brain*. Hans Huber, Toronto, pp. 225–238.
- Aertsen, A.M.H.J., Gerstein, G.L., Habib, M.K., Palm, G., 1989. Dynamics of neuronal firing correlation: modulation of 'effective connectivity'. *J. Neurophysiol.* 61, 900–917.
- Aertsen, A.M.H.J., Johannesma, P.I.M., 1980. Spectro-temporal receptive fields of auditory neurons in the grassfrog. I. Characterization of tonal and natural stimuli. *Biol. Cybern.* 38, 223–234.
- Aertsen, A.M.H.J., Johannesma, P.I.M., 1981. The spectro-temporal receptive field. A functional characterization of auditory neurons. *Biol. Cybern.* 42, 133–143.
- Aertsen, A.M.H.J., Johannesma, P.I.M., Hermes, D.J., 1980. Spectro-temporal receptive fields of auditory neurons in the grassfrog. II. Analysis of the stimulus–event relation for tonal stimuli. *Biol. Cybern.* 38, 235–248.
- Aertsen, A.M.H.J., Olders, J.H.J., Johannesma, P.I.M., 1981. Spectro-temporal receptive fields of auditory neurons in the grassfrog. III. Analysis of the stimulus–event relation for natural stimuli. *Biol. Cybern.* 39, 195–209.
- Aertsen, A.M.H.J., Smolders, J.W.T., Johannesma, P.I.M., 1979. Neural representation of the acoustic biotope: on the existence of stimulus–event relations for sensory neurons. *Biol. Cybern.* 32, 175–185.
- Aitkin, L., 1986. *The Auditory Midbrain*. Humana Press, Clifton, NJ.

- Barlow, H.B., 1961. Possible principle underlying the transformation of sensory messages. In: Rosenblith, W. (Ed.), *Sensory Communication*. MIT Press, Cambridge, MA, pp. 217–234.
- Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., Pike, B., 2000. Voice-selective areas in human auditory cortex. *Nature* 403, 309–312.
- Bieser, A., 1998. Processing of twitter-call fundamental frequencies in insula and auditory cortex of squirrel monkeys. *Exp. Brain Res.* 122, 139–148.
- Bieser, A., Müller-Preuss, P., 1996. Auditory responsive cortex in the squirrel monkey: neural responses to amplitude-modulated sounds. *Exp. Brain Res.* 108, 273–284.
- Blackburn, C.C., Sachs, M.B., 1990. The representations of the steady-state vowel sound /e/ in the discharge patterns of cat anteroventral cochlear nucleus neurons. *J. Neurophysiol.* 63, 1191–1212.
- Black, I.B., 1991. *Information in the Brain: a Molecular Perspective*. The MIT Press, Cambridge, MA.
- de Boer, E., 1967. Correlation studies applied to the frequency resolution of the cochlea. *J. Audit. Res.* 7, 209–217.
- de Boer, E., 1969. Reverse correlation. II. Initiation of nerve impulses in the inner ear. *Proc. R. Neth. Acad. Sci. C72*, 129–151.
- de Boer, E., de Jongh, H.R., 1978. On cochlear encoding: potentialities and limitations of the reverse correlation technique. *J. Acoust. Soc. Am.* 63, 115–135.
- Borst, A., Theunissen, F.E., 1999. Information theory and neural coding. *Nat. Neurosci.* 2, 947–957.
- Brosch, M., Schreiner, C.E., 1999. Correlations between neural discharges are related to receptive field properties in cat primary auditory cortex. *Eur. J. Neurosci.* 11, 3517–3530.
- Brugge, J.F., 1992. An overview of central auditory processing. In: Popper, A.N., Fay, R.R. (Eds.), *The Mammalian Auditory Pathway: Neurophysiology*. Springer Verlag, New York, pp. 1–33.
- Buchfellner, E., Leppelsack, H.-J., Klump, G.M., Häusler, U., 1989. Gap detection in the starling (*Sturnus vulgaris*). II. Coding of gaps by forebrain neurons. *J. Comp. Physiol. A* 164, 539–549.
- Buonomano, D.V., Merzenich, M.M., 1995. Temporal information transformed into a spatial code by a neural network with realistic properties. *Science* 267, 1028–1030.
- Buracas, G.T., Albright, T.D., 1999. Gauging sensory representations in the brain. *Trends Neurosci.* 22, 303–309.
- Cariani, P., 1995. As if time really mattered: temporal strategies for neural coding of sensory information. *Commun. Cogn. Artif. Intell.* 12, 157–219.
- Cariani, P., 1997. Emergence of new signal-primitives in neural systems. *Intellectica* 2, 95–143.
- Cariani, P., 1999a. Temporal coding of periodicity pitch in the auditory system: an overview. *Neural Plast.* 6, 147–172.
- Cariani, P., 1999b. Neural timing nets for auditory computation. In: Greenberg, S., Slaney, M. (Eds.), *Computational Models of Auditory Function*. IOS Press.
- Cariani, P., Delgutte, B., 1996. Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. *J. Neurophysiol.* 76, 1698–1716.
- Carr, C.E., 1993. Processing of temporal information in the brain. *Annu. Rev. Neurosci.* 16, 223–243.
- Carney, L.H., Geisler, C.D., 1986. A temporal analysis of auditory-nerve fiber responses to spoken stop consonant-vowel syllables. *J. Acoust. Soc. Am.* 79, 1896–1914.
- Cecchi, G.A., Sigman, M., Alonso, J.M., Martinez, L., Chialvo, D.R., Magnasco, M.O., 2000. Noise in neurons is message dependent. *Proc. Natl. Acad. Sci. USA* 97, 5557–5561.
- Chen, G.D., Nudrig, S.C., Narayan, S.S., Sinex, D.G., 1996. Responses of single neurons in the chinchilla inferior colliculus to consonant-vowel syllables differing in voice onset time. *Audit. Neurosci.* 3, 179–198.
- Christini, D.J., Collins, J.J., 1995. Controlling nonchaotic neuronal noise using chaos control techniques. *Phys. Rev. Lett.* 75, 2782–2785.
- Clarke, S., de Ribaupierre, F., Rouiller, E.M., de Ribaupierre, Y., 1993. Several neuronal and axonal types form long intrinsic connections in the cat primary auditory cortical field (AI). *Anat. Embryol.* 188, 117–138.
- Cohen, Y.E., Knudsen, E.I., 1999. Maps versus clusters: different representations of auditory space in the midbrain and forebrain. *Trends Neurosci.* 22, 128–135.
- Conley, R.A., Keilson, S., 1995. Rate representation and discriminability of second formant frequencies for /e/-like steady-state vowels in cat auditory nerve. *J. Acoust. Soc. Am.* 98, 3223–3234.
- Covey, E., 2000. Neural population coding and auditory temporal pattern analysis. *Physiol. Behav.* 69, 211–220.
- Covey, E., Casseday, J.H., 1999. Timing in the auditory system of the bat. *Annu. Rev. Physiol.* 61, 457–476.
- Creutzfeldt, O., Hellweg, F.-C., Schreiner, C., 1980. Thalamocortical transformation of responses to complex auditory stimuli. *Exp. Brain Res.* 39, 87–104.
- deCharms, C.R., Blake, D.T., Merzenich, M.M., 1998. Optimizing sound features for cortical neurons. *Science* 280, 1439–1443.
- deCharms, R.C., Merzenich, M.M., 1996. Primary cortical representation of sounds by the coordination of action-potential timing. *Nature* 381, 610–613.
- deCharms, R.C., Zador, A., 2000. Neural representation and the cortical code. *Annu. Rev. Neurosci.* 23, 613–647.
- Delgutte, B., 1984. Speech coding in the auditory nerve: II. processing schemes for vowel-like sounds. *J. Acoust. Soc. Am.* 75, 879–886.
- Delgutte, B., Hammond, B.M., Cariani, P.A., 1998. Neural coding of the temporal envelope of speech: relation to modulation transfer functions. In: Palmer A.R., Rees, A., Summerfield, A.Q., Meddis, R. (Eds.), *Psychophysical and Physiological Advances in Hearing*. Whurr Publishing, London, pp. 595–603.
- Delgutte, B., Kiang, N.Y.S., 1984a. Speech coding in the auditory nerve: I. vowel-like sounds. *J. Acoust. Soc. Am.* 75, 866–878.
- Delgutte, B., Kiang, N.Y.S., 1984b. Speech coding in the auditory nerve: III. voiceless fricative consonants. *J. Acoust. Soc. Am.* 75, 887–896.
- Delgutte, B., Kiang, N.Y.S., 1984c. Speech coding in the auditory nerve: IV. sounds with consonant-like dynamic characteristics. *J. Acoust. Soc. Am.* 75, 897–907.
- Deng, L., Geisler, C.D., 1987. Responses of auditory-nerve fibers to nasal consonant-vowel syllables. *J. Acoust. Soc. Am.* 82, 1977–1988.
- Dinse, H.R., Recanzone, G.H., Merzenich, M.M., 1993. Alterations in correlated activity parallel ICMS-induced representational plasticity. *NeuroReport* 68, 173–176.
- Doupe, A.J., Kuhl, P.K., 1999. Birdsong and human speech. *Annu. Rev. Neurosci.* 22, 567–631.
- Edelman, G.M., 1987. *Neural Darwinism: The Theory of Neuronal Group Selection*. Basic Books, New York.
- Eggermont, J.J., 1975. Cochlear adaptation: A theoretical description. *Biol. Cybern.* 19, 181–190.
- Eggermont, J.J., 1985. Peripheral auditory adaptation and fatigue: a model oriented review. *Hear. Res.* 18, 57–71.
- Eggermont, J.J., 1989. Coding of free field intensity in the auditory midbrain of the leopard frog. I. Results for tonal stimuli. *Hear. Res.* 40, 147–166.
- Eggermont, J.J., 1990. Temporal modulation transfer functions for single neurons in the auditory midbrain of the leopard frog. Intensity and carrier-frequency dependence. *Hear. Res.* 43, 181–198.
- Eggermont, J.J., 1991. Rate and synchronization measures of periodicity coding in cat primary auditory cortex. *Hear. Res.* 56, 153–167.

- Eggermont, J.J., 1992a. Neural interaction in cat primary auditory cortex. Dependence on recording depth, electrode separation and age. *J. Neurophysiol.* 68, 1216–1228.
- Eggermont, J.J., 1992b. Stimulus induced and spontaneous rhythmic firing of single units in cat primary auditory cortex. *Hear. Res.* 61, 1–11.
- Eggermont, J.J., 1992c. Stimulus induced and spontaneous rhythmic firing of single units in cat primary auditory cortex. *Hear. Res.* 61, 1–11.
- Eggermont, J.J., 1993a. Differential effects of age on click-rate and amplitude modulation-frequency coding in primary auditory cortex of the cat. *Hear. Res.* 65, 175–192.
- Eggermont, J.J., 1993b. Functional aspects of synchrony and correlation in the auditory nervous system. *Concepts Neurosci.* 2, 105–129.
- Eggermont, J.J., 1994a. Temporal modulation transfer functions for AM and FM stimuli in cat auditory cortex. Effects of carrier type, modulating waveform and intensity. *Hear. Res.* 74, 51–66.
- Eggermont, J.J., 1994b. Neural interaction in cat primary auditory cortex II. Effects of sound stimulation. *J. Neurophysiol.* 71, 246–270.
- Eggermont, J.J., 1995a. Representation of a voice onset time continuum in primary auditory cortex of the cat. *J. Acoust. Soc. Am.* 98, 911–920.
- Eggermont, J.J., 1995b. Neural correlates of gap detection and auditory fusion in cat auditory cortex. *NeuroReport* 6, 1645–1648.
- Eggermont, J.J., 1996. How homogeneous is cat primary auditory cortex? Evidence from simultaneous single-unit recordings. *Audit. Neurosci.* 2, 76–96.
- Eggermont, J.J., 1997a. Firing rate and firing synchrony distinguish dynamic from steady state sound. *NeuroReport* 8, 2709–2713.
- Eggermont, J.J., 1997. Representation of amplitude modulated sounds in two fields in auditory cortex of the cat. In: Syka, J. (Ed.), *Acoustical Signal Processing in the Central Auditory System*. Plenum Press, New York, pp. 303–319.
- Eggermont, J.J., 1998a. Azimuth coding in primary auditory cortex of the cat. II. Relative latency and interspike interval representation. *J. Neurophysiol.* 80, 2151–2161.
- Eggermont, J.J., 1998b. Is there a neural code? *Neurosci. Biobehav. Rev.* 22, 355–370.
- Eggermont, J.J., 1998c. Representation of spectral and temporal sound features in three cortical fields of the cat. Similarities outweigh differences. *J. Neurophysiol.* 80, 2743–2764.
- Eggermont, J.J., 1999. Neural correlates of gap detection in three auditory cortical fields in the cat. *J. Neurophysiol.* 81, 2570–2581.
- Eggermont, J.J., 2000a. Sound induced correlation of neural activity between and within three auditory cortical areas. *J. Neurophysiol.* 83, 2708–2722.
- Eggermont, J.J., 2000b. Neural responses in primary auditory cortex mimic psychophysical, across frequency-channel, gap-detection thresholds. *J. Neurophysiol.* 84, 1453–1463.
- Eggermont, J.J., 2001. Temporal modulation transfer functions in cat primary auditory cortex: dependence on carrier, modulating waveform and type of neural response. ARO Midwinter Meeting, Abstracts 24.
- Eggermont, J.J., Aertsen, A.M.H.J., Hermes, D.J., Johannesma, P.I.M., 1981. Spectro-temporal characterization of auditory neurons: redundant or necessary. *Hear. Res.* 5, 109–121.
- Eggermont, J.J., Aertsen, A.M.H.J., Johannesma, P.I.M., 1983a. Prediction of the responses of auditory neurons in the midbrain of the grassfrog based on the spectro-temporal receptive field. *Hear. Res.* 10, 191–201.
- Eggermont, J.J., Johannesma, P.I.M., Aertsen, A.M.H.J., 1983b. Reverse correlation methods in auditory research. *Q. Rev. Biophys.* 16, 341–414.
- Eggermont, J.J., Komiya, H., 2000. Moderate noise trauma in juvenile cats results in profound cortical topographic map changes in adulthood. *Hear. Res.* 142, 89–101.
- Eggermont, J.J., Schmidt, P.H., 1990. The auditory brainstem response. In: Colon, E.J., Visser, S.L. (Eds.), *Evoked Potential Manual*. Kluwer Academic, pp. 41–77.
- Eggermont, J.J., Smith, G.M., 1995. Synchrony between single-unit activity and local field potentials in relation to periodicity coding in primary auditory cortex. *J. Neurophysiol.* 73, 227–245.
- Ehret, G., 1997. The auditory cortex. *J. Comp. Physiol. A* 181, 547–557.
- Ehret, G., Merzenich, M.M., 1988a. Neuronal discharge rate is unsuitable for encoding sound intensity at the inferior colliculus level. *Hear. Res.* 35, 1–8.
- Ehret, G., Merzenich, M.M., 1988b. Complex sound analysis (frequency resolution, filtering and spectral integration) by single units of the inferior colliculus of the cat. *Brain Res. Rev.* 13, 139–163.
- Ehret, G., Schreiner, C.E., 2000. Regional variations of noise-induced changes in operating range in cat AI. *Hear. Res.* 141, 107–116.
- Epping, W.J.M., Eggermont, J.J., 1985. Single-unit characteristics in the auditory midbrain of the immobilized grassfrog. *Hear. Res.* 18, 223–243.
- Epping, W.J.M., Eggermont, J.J., 1986. Sensitivity of neurons in the auditory midbrain of the grassfrog to temporal characteristics of sound. II. Stimulation with amplitude modulated sound. *Hear. Res.* 24, 55–72.
- Erickson, R.P., 1974. Parallel ‘population’ neural coding in feature extraction. In: Schmitt, F.O., Worden, F.G. (Eds.), *The Neurosciences. Third Study Program*. MIT Press, Cambridge, MA, pp. 155–169.
- Evans, E.F., 1992. Auditory processing of complex sounds: an overview. *Phil. Trans. R. Soc. B* 336, 295–306.
- Fassel, R., Kohlrausch, A., 1995. Modulation detection as a function of carrier frequency and level. *IPO Annu. Prog. Rep.* 30, 21–29.
- Fitch, R.H., Miller, S., Tallal, P., 1997. Neurobiology of speech perception. *Annu. Rev. Neurosci.* 20, 331–353.
- Fitzpatrick, D.C., Batra, R., Stanford, T.R., Kuwada, S., 1997. A neuronal population code for sound localization. *Nature* 388, 871–874.
- Frank-Kamenetskii, M.D., 1997. *Unraveling DNA*. Addison-Wesley, Reading, MA.
- Frisina, R.D., Smith, R.L., Chamberlain, S.G., 1990. Encoding of amplitude modulation in the gerbil cochlear nucleus: I. A hierarchy of enhancement. *Hear. Res.* 44, 99–122.
- Furukawa, S., Xu, L., Middlebrooks, J.C., 2000. Coding of sound-source location by ensembles of cortical neurons. *J. Neurosci.* 20, 1216–1228.
- Gaese, B.H., Ostwald, J., 1995. Temporal coding of amplitude and frequency modulation in the rat auditory cortex. *Eur. J. Neurosci.* 7, 438–450.
- Gardner, R.B., Wilson, J.P., 1979. Evidence for direction-specific channels in the processing of frequency modulation. *J. Acoust. Soc. Am.* 66, 704–709.
- Gehr, D.D., Komiya, H., Eggermont, J.J., 2000. Neuronal responses in cat primary auditory cortex to natural and altered species-specific calls. *Hear. Res.* 150, 27–42.
- Geisler, C.D., 1988. Representation of speech sounds in the auditory nerve. *J. Phon.* 16, 19–35.
- Geisler, C.D., Gamble, T., 1989. Responses of ‘high-spontaneous’ auditory nerve fibers to consonant-vowel syllables in noise. *J. Acoust. Soc. Am.* 85, 1639–1652.
- Gielen, C.C.A.M., Hesselmann, G.H.F.M., Johannesma, P.I.M., 1988. Sensory interpretation of neural activity patterns. *Math. Biosci.* 88, 15–35.
- van Gisbergen, J.A.M., Grashuis, J.L., Johannesma, P.I.M., Vendrik,

- A.J.H., 1975. Neurons in the cochlear nucleus investigated with tone and noise stimuli. *Exp. Brain Res.* 23, 387–406.
- Graybiel, A.M., 1973. The thalamo-cortical projection of the so-called posterior nuclear group: a study with degeneration methods in the cat. *Brain Res.* 49, 229–244.
- Grothe, B., 2000. The evolution of temporal processing in the medial superior olive, an auditory brainstem structure. *Prog. Neurobiol.* 61, 581–610.
- Guinan, J.J. Jr., 1996. Physiology of olivocochlear efferents. In: Dallos, P., Popper, A.N., Fay, R.R. (Eds.), *The Cochlea*. Springer Verlag, New York, pp. 435–502.
- Halpern, B., 2000. Sensory coding, decoding, and representations: unnecessary and troublesome constructs. *Physiol. Behav.* 69, 115–118.
- Harel, N., Mori, N., Sawada, S., Mount, R.J., Harrison, R.V., 2000. Three distinct auditory areas of cortex (AI, AII, and AAF) defined by optical imaging of intrinsic signals. *Neuroimage* 11, 302–312.
- Harrison, R.V., Nagasawa, A., Smith, D.W., Stanton, S., Mount, R.J., 1991. Reorganization of auditory cortex after neonatal high frequency cochlear hearing loss. *Hear. Res.* 54, 11–19.
- He, J., 1997. Modulatory effects of regional cortical activation on the onset responses of the cat medial geniculate neurons. *J. Neurophysiol.* 77, 896–908.
- He, J., Hashikawa, T., 1998. Connections of the dorsal zone of cat auditory cortex. *J. Comp. Neurol.* 400, 334–348.
- He, J., Hashikawa, T., Ojima, H., Kinouchi, Y., 1997. Temporal integration and duration tuning in the dorsal zone of cat auditory cortex. *J. Neurosci.* 17, 2615–2625.
- Heil, P., 1997a. Auditory cortical onset responses revisited. I. First-spike timing. *J. Neurophysiol.* 77, 2616–2641.
- Heil, P., 1997b. Auditory cortical onset responses revisited. II. Response strength. *J. Neurophysiol.* 77, 2642–2660.
- Heil, P., Irvine, D.R.F., 1997. First-spike timing of auditory-nerve fibers and comparison with auditory cortex. *J. Neurophysiol.* 78, 2438–2454.
- Heil, P., Irvine, D.R.F., 1998a. Functional specialization in auditory cortex: responses to frequency-modulated stimuli in the cat's posterior auditory field. *J. Neurophysiol.* 79, 3041–3059.
- Heil, P., Irvine, D.R.F., 1998b. The posterior field P of the cat auditory cortex: coding of envelope transients. *Cereb. Cortex* 8, 125–141.
- Heil, P., Rajan, R., Irvine, R.F., 1992a. Sensitivity of neurons in cat primary auditory cortex to tones and frequency-modulated stimuli. I. Effects of variation of stimulus parameters. *Hear. Res.* 63, 108–134.
- Heil, P., Rajan, R., Irvine, R.F., 1992b. Sensitivity of neurons in cat primary auditory cortex to tones and frequency-modulated stimuli. II. Organization of response properties along the 'isofrequency' dimension. *Hear. Res.* 63, 135–156.
- Hermes, D.J., Aertsen, A.M.H.J., Johannesma, P.I.M., Eggermont, J.J., 1981. Spectro-temporal characteristics of single units in the auditory midbrain of the lightly anaesthetized grass frog (*Rana temporaria* L.) investigated with noise stimuli. *Hear. Res.* 5, 147–179.
- Hernandez, A., Zainos, A., Romo, R., 2000. Neuronal correlates of sensory discrimination in the somatosensory cortex. *Proc. Natl. Acad. Sci. USA* 97, 6191–6196.
- Hesselmans, G.H., Johannesma, P.I., 1989. Spectrotemporal interpretation of activity patterns of auditory neurons. *Math. Biosci.* 93, 31–51.
- Horikawa, J., Nasu, M., Taniguchi, I., 1998. Optical recording of responses to frequency-modulated sounds in the auditory cortex. *NeuroReport* 9, 799–802.
- Horst, J.W., Javel, E., Farley, G.R., 1986. Coding of spectral fine structure in the auditory nerve. I. Fourier analysis of period and interspike interval histograms. *J. Acoust. Soc. Am.* 79, 398–416.
- Hose, B., Langner, G., Scheich, H., 1987. Topographic representation of periodicities in the forebrain of the mynah bird: one map for pitch and rhythm? *Brain Res.* 422, 367–373.
- Huang, C.L., Winer, J.A., 2000. Auditory thalamocortical projections in the cat: laminar and areal patterns of input. *J. Comp. Neurol.* 427, 302–331.
- Hudspeth, A.J., 1989. How the ear's works work. *Nature* 341, 397–404.
- Hudspeth, A.J., 1999. How hearing happens. *Neuron* 19, 947–950.
- Hyde, P.S., Knudsen, E.I., 2000. Topographic projection from the optic tectum to the auditory space map in the inferior colliculus of the barn owl. *J. Comp. Neurol.* 421, 146–160.
- Javel, E., 1980. Coding of AM tones in the chinchilla auditory nerve: implications for the pitch of complex tones. *J. Acoust. Soc. Am.* 68, 133–146.
- Javel, E., McGee, J.A., Horst, J.W., Farley, G.R., 1988. Temporal mechanisms in auditory stimulus coding. In: Edelman, G.M., Gall, W.E., Cowan, W.M. (Eds.), *Auditory Function. Neurobiological Bases of Hearing*. J. Wiley and Sons, New York, pp. 515–558.
- Jen, P.H.-S., Chen, Q.C., Sun, X.D., 1998. Corticofugal regulation of auditory sensitivity in the bat inferior colliculus. *J. Comp. Physiol. A* 183, 683–697.
- Johannesma, P.I.M., 1981. Neural representation of sensory stimuli and sensory interpretation of neural activity. In: Székely, G., Labos, E., Damjanovitch, S. (Eds.), *Adv. Physiol. Sci.* 30, Neural Communication and Control. Pergamon Press, London, pp. 103–125.
- Johannesma, P.I.M., Aertsen, A., van den Boogaard, H., Eggermont, J.J., Epping, W., 1986. From synchrony to harmony: ideas on the function of neural assemblies and on the interpretation of neural synchrony. In: Palm, G., Aertsen, A. (Eds.), *Brain Theory*. Springer Verlag, Berlin, pp. 25–47.
- Johnson, K.O., 1980. Sensory discrimination: neural processes preceding discrimination decision. *J. Neurophysiol.* 43, 1793–1815.
- Johnson, K.O., 2000. Neural coding. *Neuron* 26, 563–566.
- Joris, P.X., 1996. Envelope coding in the lateral superior olive. II. Characteristic delays and comparison with responses in the medial superior olive. *J. Neurophysiol.* 76, 2137–2156.
- Joris, P.X., Yin, T.C.T., 1992. Responses to amplitude-modulated tones in the auditory nerve of the cat. *J. Acoust. Soc. Am.* 91, 215–232.
- Joris, P.X., Yin, T.C.T., 1998. Envelope coding in the lateral superior olive. III. Comparison with afferent pathways. *J. Neurophysiol.* 79, 253–269.
- Kaas, J.H., 1987. The organization of neocortex in mammals: implications for theories of brain function. *Annu. Rev. Psychol.* 38, 129–151.
- Kaas, J.H., Hackett, T.A., 2000. Subdivisions of auditory cortex and processing streams in primates. *Proc. Natl. Acad. Sci. USA* 97, 11793–11799.
- Kay, R.H., 1982. Hearing of modulation in sounds. *Physiol. Rev.* 62, 894–975.
- Keilson, S.E., Richards, V.M., Wyman, B.T., Young, E.D., 1997. The representation of concurrent vowels in the cat anesthetized ventral cochlear nucleus: evidence for a periodicity-tagged spectral representation. *J. Acoust. Soc. Am.* 102, 1056–1071.
- Keller, C.H., Takahashi, T.T., 2000. Representation of temporal features of complex sounds by the discharge patterns of neurons in the owl's inferior colliculus. *J. Neurophysiol.* 84, 2638–2650.
- Kiang, N.Y.S., Watanabe, T., Thomas, C., Clark, L.F., 1965. Discharge patterns of single fibers in the cat's auditory nerve. MIT Press, Cambridge, MA.

- Kilgaard, M.P., Merzenich, M.M., 1999. Distributed representation of spectral and temporal information in rat primary auditory cortex. *Hear. Res.* 134, 16–28.
- Kim, D.O., Sirianni, J.G., Chang, S.O., 1990. Responses of DCN-PVCN neurons and auditory nerve fibers in unanesthetized decerebrated cats to AM and pure tones: analysis with autocorrelation/power-spectrum. *Hear. Res.* 45, 95–113.
- Kim, P.J., Young, E.D., 1994. Comparative analysis of spectro-temporal receptive fields, reverse correlation functions and frequency tuning curves of auditory nerve fibers. *J. Acoust. Soc. Am.* 95, 410–422.
- Klump, G.M., Gleich, O., 1991. Gap detection in the European starling (*Sturnus vulgaris*). III. Processing in the peripheral auditory system. *J. Comp. Physiol. A* 168, 469–476.
- Knight, P., 1977. Representation of the cochlea within the anterior auditory field (AAF) of the cat. *Brain Res.* 130, 447–467.
- Knudsen, E.I., du Lac, S., Esterly, S.D., 1987. Computational maps in the brain. *Annu. Rev. Neurosci.* 10, 41–65.
- Koch, U., Grothe, B., 2000. Interdependence of spatial and temporal coding in the auditory midbrain. *J. Neurophysiol.* 83, 2300–2314.
- Koenderink, J.J., 1984. Simultaneous order in nervous nets from a functional standpoint. *Biol. Cybern.* 50, 35–41.
- Konishi, M., Takahashi, T.T., Wagner, H., Sullivan, W.E., Carr, C.C., 1988. Neurophysiological and anatomical substrates of sound localization in the owl. In: Edelman, G.M., Gall, W.E., Cowan, W.M. (Eds.), *Auditory Function. Neurobiological Bases of Hearing*. J. Wiley and Sons, New York, pp. 721–745.
- Köppl, C., 1997. Phase-locking to high frequencies in the auditory nerve and cochlear nucleus magnocellularis of the barn owl, *Tyto alba*. *J. Neurosci.* 17, 3312–3321.
- Kowalski, N., Depireux, D.A., Shamma, S.A., 1996. Analysis of dynamic spectra in ferret primary auditory cortex. I. Characteristics of single-unit responses to moving ripple spectra. *J. Neurophysiol.* 76, 3503–3523.
- Krishna, B.S., Semple, M.N., 2000. Auditory temporal processing: responses to sinusoidally amplitude-modulated tones in the inferior colliculus. *J. Neurophysiol.* 84, 255–273.
- Kuhl, P.K., Miller, J.D., 1978. Speech perception by the chinchilla: identification functions for synthetic VOT stimuli. *J. Acoust. Soc. Am.* 63, 905–917.
- Kuwada, S., Batra, R., 1999. Coding of sound envelopes by inhibitory rebound in neurons of the superior olivary complex in the unanesthetized rabbit. *J. Neurosci.* 19, 2273–2287.
- Langner, G., 1981. Neuronal mechanisms for pitch analysis in the time domain. *Exp. Brain Res.* 44, 450–454.
- Langner, G., 1992. Periodicity coding in the auditory system. *Hear. Res.* 60, 115–142.
- Langner, G., Schreiner, C.E., 1988. Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. *J. Neurophysiol.* 60, 1799–1822.
- Langner, G., Schreiner, C.E., 1996. Contributions of the auditory brainstem to periodicity pitch coding. In: *Advances in Speech, Hearing and Language Processing*, Vol. 3. JAI Press, pp. 447–461.
- Leppelsack, H.-J., 1978. Unit responses to species-specific sounds in the auditory forebrain center of birds. *Fed. Proc.* 37, 2336–2341.
- Le Prell, G., Sachs, M., May, B., 1996. Representation of vowel-like spectra by discharge rate responses of individual auditory nerve fibers. *Audit. Neurosci.* 2, 275–288.
- Lewicki, M.S., Arthur, B., 1996. Hierarchical organization of auditory temporal context sensitivity. *J. Neurosci.* 16, 6987–6998.
- Liberman, A.M., Cooper, F.S., Shankweiler, D.P., Studdert-Kennedy, M., 1967. Perception of the speech code. *Psychol. Rev.* 74, 431–461.
- Liberman, M.C., 1978. Auditory-nerve responses from cats raised in a low-noise chamber. *J. Acoust. Soc. Am.* 63, 442–455.
- Liberman, M.C., 1988. Response properties of cochlear efferent neurons: monaural vs. binaural stimulation and the effects of noise. *J. Neurophysiol.* 60, 1779–1798.
- Liberman, M.C., 1991. Central projections of auditory-nerve fibers of differing spontaneous rate. I. Anteroventral cochlear nucleus. *J. Comp. Neurol.* 313, 240–258.
- Liberman, M.C., 1993. Central projections of auditory-nerve fibers of differing spontaneous rate. II. Posteroventral and dorsal cochlear nuclei. *J. Comp. Neurol.* 327, 17–36.
- Lisman, J.E., 1997. Bursts as a unit of neural information: making unreliable synapses reliable. *Trends Neurosci.* 20, 38–43.
- Loeb, G.E., White, M.W., Merzenich, M.M., 1983. Spatial cross-correlation. A proposed method for acoustic pitch perception. *Biol. Cybern.* 47, 149–163.
- Lu, T., Wang, X., 2000. Temporal discharge patterns evoked by rapid sequences of wide and narrowband clicks in primary auditory cortex of cat. *J. Neurophysiol.* 84, 236–246.
- May, B.J., Huang, A., Le Prell, G., Hiez, R.D., 1996. Vowel formant frequency discrimination in cats: comparison of auditory nerve representations and psychophysical thresholds. *Audit. Neurosci.* 3, 135–162.
- May, B.J., Le Prell, G.S., Sachs, M.B., 1998. Vowel representations in the ventral cochlear nucleus of the cat: effects of level, background noise, and behavioral state. *J. Neurophysiol.* 79, 1755–1767.
- Mendelson, J.R., Schreiner, C.E., Sutter, M.L., Grasse, K.L., 1993. Functional topography of cat primary auditory cortex: responses to frequency-modulated sweeps. *Exp. Brain Res.* 94, 65–87.
- Merzenich, M.M., deCharms, R.C., 1996. Neural representations, experience, and change. In: Llinas, R., Churchland, P.S. (Eds.), *The Mind-Brain Continuum*. MIT Press, Cambridge, MA, pp. 61–81.
- Merzenich, M.M., Knight, P., Roth, G.L., 1975. Representation of the cochlea within primary auditory cortex in the cat. *J. Neurophysiol.* 38, 231–249.
- Middlebrooks, J.C., 1988. Auditory mechanisms underlying a neural code for space in the cat's superior colliculus. In: Edelman, G.M., Gall, W.E., Cowan, W.M. (Eds.), *Auditory Function. Neurobiological Bases of Hearing*. J. Wiley and Sons, New York, pp. 431–455.
- Middlebrooks, J.C., Clock, A.E., Xu, L., Green, D.M., 1994. A panoramic code for sound location by cortical neurons. *Science* 264, 842–844.
- Miller, M.I., Sachs, M.B., 1984. Representation of voice pitch in the discharge patterns of auditory-nerve fibers. *Hear. Res.* 14, 257–279.
- Miller, R.L., Calhoun, B.M., Young, E.D., 1999. Discriminability of vowel representations in cat auditory-nerve fibers after acoustic trauma. *J. Acoust. Soc. Am.* 105, 311–325.
- Møller, A.R., 1972. Coding of amplitude and frequency modulated sounds in the cochlear nucleus of the rat. *Acta Physiol. Scand.* 86, 223–238.
- Møller, A.R., 1973. Statistical evaluation of the dynamic properties of cochlear nucleus units using stimuli modulated with pseudorandom noise. *Brain Res.* 57, 443–456.
- Møller, A.R., 1974. Coding of sounds with rapidly varying spectrum in the cochlear nucleus. *J. Acoust. Soc. Am.* 55, 631–640.
- Møller, A.R., 1976a. Dynamic properties of primary auditory fibers compared with cells in the cochlear nucleus. *Acta Physiol. Scand.* 98, 157–167.
- Møller, A.R., 1976b. Dynamic properties of the responses of single neurons in the cochlear nucleus of the rat. *J. Physiol.* 259, 63–82.
- Møller, A.R., Rees, A., 1986. Dynamic properties of the responses of single neurons in the inferior colliculus of the rat. *Hear. Res.* 24, 203–215.
- Moody, D.B., 1994. Detection and discrimination of amplitude-

- modulated signals by macaque monkeys. *J. Acoust. Soc. Am.* 95, 3499–3510.
- Morse, P.A., Snowden, C.T., 1975. An investigation of categorical speech perception by rhesus monkeys. *Percept. Psychophys.* 17, 9–16.
- Mountcastle, V.B., 1975. The view from within: pathways to the study of perception. *Johns Hopkins Med. J.* 136, 109–131.
- Mullennix, J.W., Pisoni, D.B., 1989. Speech perception: analysis of biologically significant signals. In: Dooling, R.J., Hulse, S.H. (Eds.), *The Comparative Psychology of Audition: Perceiving Complex Sounds*. Lawrence Erlbaum, Hillsdale, NJ, pp. 97–128.
- Müller, C.M., Leppelsack, H.-J., 1985. Feature extraction and tonotopic organization in the avian auditory forebrain. *Exp. Brain Res.* 59, 587–599.
- Müller-Preuss, P., Flachskamm, C., Bieser, A., 1994. Neural encoding of amplitude modulation within the auditory midbrain of squirrel monkeys. *Hear. Res.* 80, 197–208.
- Mumford, D., 1991. On the computational architecture of the neocortex. I. The role of the thalamo-cortical loop. *Biol. Cybern.* 65, 135–145.
- Mumford, D., 1992. On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biol. Cybern.* 66, 241–251.
- Nelken, I., Rotman, Y., Yosef, O.B., 1999. Responses of auditory-cortex neurons to structural features of natural sounds. *Nature* 397, 154–157.
- Nelken, I., Versnel, H., 2000. Responses to linear and logarithmic frequency-modulated sweeps in ferret primary auditory cortex. *Eur. J. Neurosci.* 12, 549–562.
- Newman, J.D., Wollberg, Z., 1973. Multiple coding of species-specific vocalizations in the auditory cortex of squirrel monkeys. *Brain Res.* 54, 287–304.
- Oertel, D., 1999. The role of timing in the brainstem auditory nuclei of vertebrates. *Ann. Rev. Physiol.* 61, 497–519.
- Ohl, F.W., Scheich, H., 1997. Orderly cortical representation of vowels along for formant interaction. *Proc. Natl. Acad. Sci. USA* 94, 9440–9444.
- Ohlemiller, K.K., Kanwal, J.S., Butman, J.A., Suga, N., 1994. Stimulus design for auditory neuroethology: synthesis and manipulation of complex communication sounds. *Audit. Neurosci.* 1, 19–37.
- Oliver, D.L., Huerta, M.F., 1992. Inferior and superior colliculi. In: Webster, D.B., Popper, A.N., Fay, R.R. (Eds.), *The Mammalian Auditory Pathway: Neuroanatomy*. Springer Verlag, New York, pp. 168–221.
- Palmer, A.R., 1982. Encoding of rapid amplitude fluctuations by cochlear nerve fibres in the guinea pig. *Arch. Otorhinolaryngol.* 236, 197–202.
- Palmer, A.R., Winter, I.M., Darwin, C.J., 1986. The representation of steady-state vowels sounds in the temporal discharge patterns of the guinea pig cochlear nerve and primarylike cochlear nucleus neurons. *J. Acoust. Soc. Am.* 79, 100–113.
- Pena, J.L., Viète, S., Albeck, Y., Konishi, M., 1996. Tolerance to sound intensity of binaural coincidence detection in the nucleus laminaris of the owl. *J. Neurosci.* 16, 7046–7054.
- Perkel, D.H., Bullock, T.H., 1969. Neural coding. In: Schmitt, F.O., Melenchuk, T., Quarton, G.C., Adelman, G. (Eds.), *Neurosciences Research Symposium Summaries*, Vol. 3. MIT Press, Cambridge, MA, pp. 405–528.
- Phillips, D.P., 1995. Central auditory processing: a view from auditory neuroscience. *Am. J. Otol.* 16, 338–352.
- Phillips, D.P., 1999. Auditory gap detection, perceptual channels, and temporal resolution in speech perception. *J. Am. Acad. Audiol.* 10, 343–354.
- Phillips, D.P., Hall, S.E., Hollett, J.L., 1989. Repetition rate and signal level effects on neuronal responses to brief tone pulses in cat auditory cortex. *J. Acoust. Soc. Am.* 85, 2537–2549.
- Phillips, D.P., Irvine, D.R.F., 1982. Properties of single neurons in the anterior auditory field (AAF) of cat cerebral cortex. *Brain Res.* 248, 237–244.
- Phillips, D.P., Mendelson, J.R., Cynader, M.S., Douglas, R.M., 1985. Responses of single neurons in cat auditory cortex to time-varying stimuli: frequency-modulated tones of narrow excursion. *Exp. Brain Res.* 58, 443–454.
- Phillips, D.P., Orman, S.S., 1984. Responses of single neurons in posterior field of cat auditory cortex to tonal stimulation. *J. Neurophysiol.* 51, 147–163.
- Phillips, D.P., Semple, M.N., Calford, M.B., Kitzes, L.M., 1994. Level-dependent representation of stimulus frequency in cat primary auditory cortex. *Exp. Brain Res.* 102, 210–226.
- Phillips, D.P., Semple, M.N., Kitzes, L.M., 1995. Factors shaping the tone level sensitivity of single neurons in posterior field of cat auditory cortex. *J. Neurophysiol.* 73, 674–686.
- Plomp, R., 1976. *Aspects of Tone Sensation*. Academic Press, New York.
- Ponton, C.W., Eggermont, J.J., Don, M., Waring, M.D., Kwong, B., Cunningham, J., Trautwein, P., 2000. Maturation of mismatch negativity: effects of profound deafness and cochlear implant use. *Audiol. Neurotol.* 5, 167–185.
- Poon, P.W.F., Chen, X., Cheung, Y.M., 1992. Differences in FM response correlate with morphology of neurons in the rat inferior colliculus. *Exp. Brain Res.* 91, 94–104.
- Preuss, A., Müller-Preuss, P., 1990. Processing of amplitude modulated sounds in the medial geniculate body of squirrel monkeys. *Exp. Brain Res.* 79, 207–211.
- Rajan, R., Irvine, D.R.F., Wise, L.Z., Heil, P., 1993. Effect of unilateral partial cochlear lesions in adult cats on the representation of lesioned and unlesioned cochleas in primary auditory cortex. *J. Comp. Neurol.* 338, 17–49.
- Rauschecker, J.P., 1998. Parallel processing in the auditory cortex of primates. *Audiol. Neurotol.* 3, 86–103.
- Rauschecker, J.P., Tian, B., 2000. Mechanisms and streams for processing of ‘what’ and ‘where’ in auditory cortex. *Proc. Natl. Acad. Sci. USA* 97, 11800–11806.
- Reale, R.A., Imig, T.J., 1980. Tonotopic organization in auditory cortex of the cat. *J. Comp. Neurol.* 192, 265–291.
- Recanzone, G.H., Guard, D.C., Phan, M.L., 2000a. Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey. *J. Neurophysiol.* 83, 2315–2331.
- Recanzone, G.H., Guard, D.C., Phan, M.L., Su, T.-I.K., 2000b. Correlation between the activity of single cortical neurons and sound-localization behavior in the macaque monkey. *J. Neurophysiol.* 83, 2723–2739.
- Reccio, A., Rhode, W.S., 2000. Representation of vowel stimuli in the ventral cochlear nucleus of the chinchilla. *Hear. Res.* 146, 167–184.
- Rees, A., Möller, A.R., 1983. Responses of neurons in the inferior colliculus of the rat to AM and FM tones. *Hear. Res.* 10, 310–330.
- Rees, A., Möller, A.R., 1987. Stimulus properties influencing the responses of inferior colliculus neurons to amplitude-modulated sounds. *Hear. Res.* 27, 129–143.
- Regan, D., Tansley, B.W., 1979. Selective adaptation to frequency-modulated tones: evidence for an information-processing channel selectively sensitive to frequency changes. *J. Acoust. Soc. Am.* 65, 1249–1257.
- Rhode, W.S., 1991. Physiological-morphological properties of the cochlear nucleus. In: Altshuler, R.A., Bobbin, R.P., Clopton, B.M., Hoffman, D.W. (Eds.), *Neurobiology of Hearing: the Central Auditory System*. Raven Press, New York, pp. 47–77.
- Richards, D.G., Wiley, R.H., 1980. Reverberations and amplitude

- fluctuations in the propagation of sound in a forest: implication for animal communication. *Am. Nat.* 115, 381–399.
- Rieke, F., Bodnar, D.A., Bialek, W., 1995. Naturalistic stimuli increase the rate and efficiency of information transmission by primary auditory afferents. *Proc. R. Soc. Lond. B* 262, 259–265.
- Rieke, F., Warland, D., de Ruyter van Steveninck, R., Bialek, W., 1997. *Spikes. Exploring the Neural Code*. MIT Press, Cambridge, MA.
- Romanski, L.M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P.S., Rauschecker, J.P., 1999. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat. Neurosci.* 2, 1131–1136.
- Rouiller, E., de Ribaupierre, F., 1982. Neurons sensitive to narrow ranges of repetitive acoustic transients in the medial geniculate body of the cat. *Exp. Brain Res.* 48, 323–326.
- Ruderman, D.L., Bialek, W., 1994. Statistics of natural images: scaling in the woods. *Phys. Rev. Lett.* 73, 814–817.
- Ryugo, D., 1992. The auditory nerve: peripheral innervation, cell body morphology, and central projections. In: Webster, D.B., Popper, A.N., Fay, R.R. (Eds.), *The Mammalian Auditory Pathway: Neuroanatomy*. Springer Verlag, New York, pp. 23–65.
- Saberi, K., Hafner, E.R., 1995. A common neural code for frequency- and amplitude-modulated sounds. *Nature* 374, 537–539.
- Sachs, M.B., 1984. *Speech encoding in the auditory nerve*. In: Berlin, C.I. (Ed.), *Hearing Science*. College-Hill Press, San Diego, CA, pp. 263–307.
- Sachs, M.B., Voigt, H.F., Young, E.D., 1983. Auditory nerve representation of vowels in background noise. *J. Neurophysiol.* 50, 27–45.
- Sachs, M.B., Young, E.D., 1979. Encoding of steady-state vowels in the auditory nerve: representation in terms of discharge rate. *J. Acoust. Soc. Am.* 66, 470–479.
- Salvi, R.J., Giraudi, D.M., Henderson, D., Hamernik, R.P., 1982. Detection of sinusoidally amplitude modulated noise by the chinchilla. *J. Acoust. Soc. Am.* 71, 424–429.
- Schiff, S.J., Jerger, K., Duong, D.H., Chang, T., Spano, M.L., Ditto, W.L., 1994. Controlling chaos in the brain. *Nature* 370, 615–620.
- Schreiner, C.E., 1995. Order and disorder in auditory cortical maps. *Curr. Opin. Neurobiol.* 5, 484–496.
- Schreiner, C.E., Calhoun, B.M., 1994. Spectral envelope coding in cat primary auditory cortex: properties of ripple transfer functions. *Audit. Neurosci.* 1, 39–61.
- Schreiner, C.E., Cynader, M.S., 1984. Basic functional organization of secondary auditory cortical field (AII) of the cat. *J. Neurophysiol.* 51, 1284–1305.
- Schreiner, C.E., Langner, G., 1988. Periodicity coding in the inferior colliculus of the cat. II. Topographic organization. *J. Neurophysiol.* 60, 1823–1840.
- Schreiner, C.E., Langner, G., 1997. Laminar fine structure of frequency organization in auditory midbrain. *Nature* 388, 383–386.
- Schreiner, C.E., Mendelson, J.R., 1990. Functional topography of cat primary auditory cortex: distribution of integrated excitation. *J. Neurophysiol.* 64, 1442–1459.
- Schreiner, C.E., Raggio, M.W., 1996. Neuronal responses in cat primary auditory cortex to electrical cochlear stimulation. II. Repetition rate coding. *J. Neurophysiol.* 75, 1283–1300.
- Schreiner, C.E., Read, H.L., Sutter, M.L., 2000. Modular organization of frequency integration in primary auditory cortex. *Annu. Rev. Neurosci.* 23, 501–529.
- Schreiner, C.E., Sutter, M.L., 1992. Topography of excitatory bandwidth in cat primary auditory cortex: single-neuron versus multiple-neuron recordings. *J. Neurophysiol.* 68, 1487–1502.
- Schreiner, C.E., Urbas, J.V., 1986. Representation of amplitude modulation in the auditory cortex of the cat. I. The anterior auditory field (AAF). *Hear. Res.* 21, 227–241.
- Schreiner, C.E., Urbas, J.V., 1988. Representation of amplitude modulation in the auditory cortex of the cat. II. Comparison between cortical fields. *Hear. Res.* 32, 49–64.
- Schulze, H., Langner, G., 1997. Periodicity coding in the primary auditory cortex of the Mongolian gerbil (*Meriones unguiculatus*): two different coding strategies for pitch and rhythm? *J. Comp. Physiol. A* 181, 651–663.
- Seung, H.S., Sompolinsky, H., 1993. Simple models for reading neuronal population codes. *Proc. Natl. Acad. Sci. USA* 90, 10749–10753.
- Shamma, S.A., 1985a. Speech processing in the auditory system I: The representation of speech sounds in the responses of the auditory nerve. *J. Acoust. Soc. Am.* 78, 1612–1621.
- Shamma, S.A., 1985b. Speech processing in the auditory system. II: lateral inhibition and the central processing of speech evoked activity in the auditory nerve. *J. Acoust. Soc. Am.* 78, 1622–1632.
- Shamma, S.A., 1996. Auditory cortical representation of complex acoustic spectra as inferred from the ripple analysis method. *Netw. Comput. Neural Syst.* 7, 439–476.
- Shamma, S.A., Fleshman, J.W., Wiser, P.R., Versnel, H., 1993. Organization of response areas in ferret primary auditory cortex. *J. Neurophysiol.* 69, 367–383.
- Shamma, S.A., Versnel, H., Kowalski, N., 1995. Ripple analysis in the ferret primary auditory cortex. I. Response characteristics of single units to sinusoidally rippled spectra. *Audit. Neurosci.* 1, 233–254.
- Shannon, C.E., 1948. A mathematical theory of communication. *AT&T Bell L. Tech. J.* 27, 379–423.
- Shannon, R.V., Zeng, F.-G., Kamath, V., Wygonski, J., Ekelid, M., 1995. Speech recognition with primarily temporal cues. *Science* 270, 303–304.
- Silkes, S.M., Geisler, C.D., 1991. Responses of 'lower-spontaneous-rate' auditory nerve fibers to speech syllables presented in noise. I: General characteristics. *J. Acoust. Soc. Am.* 90, 3122–3139.
- Sinex, D.G., Geisler, C.D., 1981. Auditory-nerve fiber responses to frequency-modulated tones. *Hear. Res.* 4, 127–148.
- Sinex, D.G., Geisler, C.D., 1983. Responses of auditory-nerve fibers to consonant-vowel syllables. *J. Acoust. Soc. Am.* 73, 602–615.
- Sinex, D.G., McDonald, L.P., 1989. Synchronized discharge rate representation of voice onset time in the chinchilla auditory nerve. *J. Acoust. Soc. Am.* 85, 1995–2004.
- Sinex, D.G., McDonald, L.P., Mott, J.B., 1991. Neural correlates of nonmonotonic temporal acuity for voice onset time. *J. Acoust. Soc. Am.* 90, 2441–2449.
- Sinex, D.G., Narayan, S.S., 1994. Auditory-nerve fiber representation of temporal cues to voicing in word-medial stop consonants. *J. Acoust. Soc. Am.* 95, 897–903.
- Smith, P.H., Joris, P.X., Yin, T.C.T., 1993. Projections of physiologically characterized spherical bushy cell axons from the cochlear nucleus of the cat: evidence for delay lines to the medial superior olive. *J. Comp. Neurol.* 331, 245–260.
- Smith, R.L., Brachman, M.L., 1982. Adaptation in auditory nerve fibers: a revised model. *Biol. Cybern.* 44, 107–120.
- Smolders, J.W.T., Aertsen, A.M.H.J., Johannesma, P.I.M., 1979. Neural representation of the acoustic biotope. A comparison of the response of auditory neurons to tonal and natural stimuli in the cat. *Biol. Cybern.* 35, 11–20.
- Snyder, R.L., Vollmer, M., Moore, C.M., Rebscher, S.J., Leake, P.A., Beitel, R.E., 2000. Responses of inferior colliculus neurons to amplitude-modulated intracochlear electrical pulses in deaf cats. *J. Neurophysiol.* 84, 166–183.
- Sovijärvi, A.R.A., 1975. Detection of natural complex sounds by cells in the primary auditory cortex of the cat. *Acta Physiol. Scan.* 93, 318–335.
- Spangler, K.M., Warr, W.B., 1991. The descending auditory system. In: Altschuler, R.A., Bobbin, R.P., Clopton, B.M., Hoffman,

- D.W. (Eds.), *Neurobiology of Hearing the Central Auditory System*. Raven Press, New York, pp. 27–45.
- Steinmetz, P.N., Roy, A., Fitzgerald, P.J., Hsiao, S.S., Johnson, K.O., Niebur, E., 2000. Attention modulates synchronized neuronal firing in primate somatosensory cortex. *Nature* 404, 187–190.
- Steinschneider, M., Arrezo, J., Vaughan, H.G., 1982. Speech-evoked activity in the auditory radiations and cortex of the awake monkey. *Brain Res.* 252, 353–365.
- Steinschneider, M., Schroeder, C.E., Arrezo, J., Vaughan, H.G., 1990. Tonotopic features of speech evoked activity in primate auditory cortex. *Brain Res.* 519, 158–168.
- Steinschneider, M., Schroeder, C.E., Arrezo, J., Vaughan, H.G., 1994. Speech-evoked activity in primary auditory cortex: effects of voice-onset time. *Electroencephalogr. Clin. Neurophysiol.* 92, 30–43.
- Stevens, C.F., Zador, A.M., 1998. Input synchrony and the irregular firing of cortical neurons. *Nature Neurosci.* 1, 210–217.
- Suga, N., 1988. Auditory neuroethology and speech processing: complex-sound processing by combination-sensitive neurons. In: Edelman, G.M., Gall, W.E., Cowan, W.M. (Eds.), *Auditory Function. Neurobiological Bases of Hearing*. J. Wiley and Sons, New York, pp. 679–720.
- Suga, N., 1989. Principles of auditory information-processing derived from neuroethology. *J. Exp. Biol.* 146, 277–286.
- Suga, N., 1992. Philosophy and stimulus design for neuroethology of complex-sound processing. *Phil. Trans. R. Soc. Lond. B* 336, 423–428.
- Suga, N., 1994. Multi-function theory for cortical processing of auditory information: implications of single unit and lesion data for future research. *J. Comp. Physiol. A* 175, 135–144.
- Suga, N., 1996. Basic acoustic patterns and neural mechanisms shared by humans and animals for auditory perception: a neuroethologist's view. In: Ainsworth, W., Greenberg, S. (Eds.), *Auditory Basis of Speech Perception. Proceedings Keele University Workshop*, pp. 31–38.
- Suga, N., Gao, E., Zhang, Y., Ma, X., Olsen, J.F., 2000. The corticofugal system for hearing: recent progress. *Proc. Natl. Acad. Sci. USA* 97, 11807–11814.
- Sutter, M.L., Schreiner, C.E., 1991. Physiology and topography of neurons with multi-peaked tuning curves in cat primary auditory cortex. *J. Neurophysiol.* 65, 1207–1226.
- Sutter, M.L., Schreiner, C.E., McLean, M., O'Connor, K.N., Loftus, W.C., 1999. Organization of inhibitory frequency receptive fields in cat primary auditory cortex. *J. Neurophysiol.* 82, 2358–2371.
- Theunissen, F.E., Doupe, A.J., 1998. Temporal and spectral sensitivity of complex auditory neurons in the nucleus HVC of male zebra finches. *J. Neurosci.* 18, 3786–3802.
- Theunissen, F.E., Sen, K., Doupe, A.J., 2000. Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *J. Neurosci.* 20, 2315–2331.
- Theunissen, F., Miller, J.P., 1995. Temporal encoding in nervous systems: a rigorous definition. *J. Comput. Neurosci.* 2, 149–162.
- Tian, B., Rauschecker, J.P., 1994. Processing of frequency-modulated sounds in the cat's anterior auditory field. *J. Neurophysiol.* 71, 1959–1975.
- Tian, B., Rauschecker, J.P., 1998. Processing of frequency-modulated sounds in the cat's posterior auditory field. *J. Neurophysiol.* 79, 2629–2642.
- Trussell, L.O., 1997. Cellular mechanisms for preservation of timing in central auditory pathways. *Curr. Opin. Neurobiol.* 7, 487–492.
- Ts'o, D.Y., Gilbert, C.D., Wiesel, T.N., 1986. Relationships between horizontal interactions and functional architecture in the cat striate cortex as revealed by cross-correlation analysis. *J. Neurosci.* 6, 1160–1170.
- Uttal, W.R., 1969. Emerging principles of sensory coding. *Perspect. Biol. Med.* 12, 344–368.
- Vaadia, E., Haalman, I., Abeles, M., Bergman, H., Prut, Y., Slovin, H., Aertsen, A., 1995. Dynamics of neuronal interactions in monkey cortex in relation to behavioural events. *Nature* 373, 515–518.
- Vater, M., 1982. Single unit responses in cochlear nucleus of horseshoe bats to sinusoidal frequency and amplitude modulated signals. *J. Comp. Physiol.* 149, 369–388.
- Versnel, H., Kowalski, N., Shamma, S.A., 1995. Ripple analysis in ferret primary auditory cortex. III. Topographic distribution of ripple response parameters. *Audit. Neurosci.* 1, 271–285.
- Versnel, H., Shamma, S.A., 1998. Spectral-ripple representation of steady-state vowels in primary auditory cortex. *J. Acoust. Soc. Am.* 103, 2502–2514.
- Viemeister, N., 1979. Temporal modulation transfer functions based upon modulation thresholds. *J. Acoust. Soc. Am.* 66, 1364–1380.
- Villa, A.E.P., Rouiller, E.M., Simm, G.M., de Ribaupierre, Y., de Ribaupierre, F., 1991. Corticofugal modulation of the information processing in the auditory thalamus of the cat. *Exp. Brain Res.* 86, 506–517.
- Villa, A.E.P., Tetko, I.V., Dutoit, P., de Ribaupierre, Y., de Ribaupierre, F., 1999. Corticofugal modulation of functional connectivity within the auditory thalamus of rat, guinea pig and cat revealed by cooling deactivation. *J. Neurosci. Methods* 86, 161–178.
- Volkov, I.O., Galazjuk, A.V., 1991. Formation of spike response to sound tones in cat auditory cortex neurons: interaction of excitatory and inhibitory effects. *Neuroscience* 43, 307–321.
- Wakefield, G.H., Viemeister, N.F., 1984. Selective adaptation to linear frequency-modulated sweeps: evidence for direction-specific FM channels? *J. Acoust. Soc. Am.* 75, 1588–1592.
- Wallace, M.N., Kitzes, L.M., Jones, E.G., 1991. Intrinsic inter- and intralaminar connections and their relationships to the tonotopic map in cat primary auditory cortex. *Exp. Brain Res.* 86, 527–544.
- Wang, K., Shamma, S.A., 1995. Spectral shape analysis in the central auditory system. *IEEE Trans. Speech Audio Process.* 3, 382–395.
- Wang, X., 1998. What is the neural code of species-specific communication sounds in the auditory cortex of primates? In: Palmer, A.R., Rees, A., Summerfield, A.Q., Meddis, R. (Eds.), *Psychophysical and Physiological Advances in Hearing*. Whurr Publishing, London, pp. 521–528.
- Wang, X., 2000. On cortical coding of vocal communication sounds in primates. *Proc. Natl. Acad. Sci. USA* 97, 11843–11849.
- Wang, X., Merzenich, M.M., Beitel, R., Schreiner, C.E., 1995. Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. *J. Neurophysiol.* 74, 2685–2706.
- Westerman, L.A., Smith, R.L., 1984. Rapid and short-term adaptation in auditory nerve responses. *Hear. Res.* 15, 249–260.
- Winer, J.A., 1990. The functional architecture of the medial geniculate body and the primary auditory cortex. In: Webster, D.B., Popper, A.N., Fay, R.R. (Eds.), *The Mammalian Auditory Pathway: Neuroanatomy*. Springer Verlag, New York, pp. 222–409.
- Winslow, R.L., Sachs, M.B., 1987. Effect of electrical stimulation of the crossed olivocochlear bundle on auditory nerve responses to tones in noise. *J. Neurophysiol.* 57, 1002–1021.
- Winter, I.A., Palmer, A.R., 1990. Temporal response of primarylike anteroventral cochlear nucleus units to the steady-state vowel /i/. *J. Acoust. Soc. Am.* 88, 1437–1441.
- Winter, P., Funkenstein, H.H., 1973. The effect of species-specific vocalizations on the discharge of auditory cortical cells in the awake squirrel monkey (*Saimiri sciureus*). *Exp. Brain Res.* 18, 489–504.
- Worden, F.G., Galambos, R., 1972. Auditory processing of biologically significant sounds. *Neurosci. Res. Prog. Bull.*, Vol. 10, Neurosciences Research Program, Brookline, MA.

- Wu, S.H., Kelly, J.B., 1995. In vitro brain slice studies of the rat's dorsal nucleus of the lateral lemniscus. I. Membrane and synaptic response properties. *J. Neurophysiol.* 73, 780–793.
- Yan, J., Suga, N., 1996. Corticofugal modulation of time-domain processing of biosonar information in bats. *Science* 273, 1100–1103.
- Yan, J., Suga, N., 1999. Corticofugal amplification of facilitative auditory response of subcortical combination-sensitive neurons in the mustached bat. *J. Neurophysiol.* 81, 817–824.
- Yan, W., Suga, N., 1998. Corticofugal modulation of the midbrain frequency map in the bat auditory system. *Nat. Neurosci.* 1, 54–58.
- Yates, G.K., 1987. Dynamic effects in the input/output relationship of auditory nerve. *Hear. Res.* 27, 221–230.
- Yin, T.C.T., Chan, J.C.K., 1988. Neural mechanisms underlying interaural time sensitivity to tones and noise. In: Edelman, G.M., Gall, W.E., Cowan, W.M. (Eds.), *Auditory Function. Neurobiological Bases of Hearing*. J. Wiley and Sons, New York, pp. 385–430.
- Young, E., 1997. Parallel processing in the central nervous system: evidence from sensory maps. *Proc. Natl. Acad. Sci. USA* 94, 933–934.
- Young, E.D., 1998. Cochlear nucleus. In: Shepherd, G.M. (Ed.), *The Synaptic Organization of the Brain*. Oxford University Press, New York, pp. 121–157.
- Young, E.D., Sachs, M.B., 1979. Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory nerve fibers. *J. Acoust. Soc. Am.* 66, 1381–1403.
- Zhang, Y., Suga, N., Yan, J., 1997. Corticofugal modulation of frequency processing in bat auditory system. *Nature* 387, 900–903.
- Zhang, Y., Suga, N., 2000. Modulation of response and frequency tuning of thalamic and collicular neurons by cortical activation in mustached bats. *J. Neurophysiol.* 84, 325–333.
- Zheng, X.Y., Henderson, D., McFadden, S.L., Ding, D.L., Salvi, R.J., 1999. Auditory nerve fiber responses following chronic cochlear de-efferentation. *J. Comp. Neurol.* 406, 72–86.
- Zwicker, H., Fastl, H., 1990. *Psychoacoustics. Facts and Models*. Springer Verlag, Berlin.