Hearing Research, 57 (1992) 231-244

HEARES 01677

Setting complex tasks to single units in the avian auditory forebrain. II: Do we really need natural stimuli to describe neuronal response characteristics?

Michael Schäfer, Rudolf Rübsamen, Gerd Joachim Dörrscheidt and Matthias Knipschild

Lehrstuhl für Allgemeine Zoologie und Neurobiologie, Ruhr-Universitat Bochum, Bochum, F.R.G.

(Received 10 March 1991; accepted 1 September 1991)

The response characteristic of auditory forebrain neurons in the European starling was established both with artificial stimuli (AS) and a conspecific territorial song as a natural stimulus (NS1). Applying experimenter-centred statistical methods for response detection and for scaling response strength, and spike-triggered analyses for the delimitation of the key sound parameters (spectrotemporal receptive field STRF, Aertsen et al. 1980) the study aimed at disclosing differences in the processing of the two stimulus classes, AS and NS.

With the STRF as reference, we find congruence (1) in the best frequency with those determined under sweep and bandpass noise stimulation, (2) in response latency, and (3) in response-intensity dependence, further similarity in the overall frequency characteristic.

Partitioning the song into 42 acoustically defined segments allowed to further delimit the response criteria under natural stimulation. They are easily understood from the AS response characteristics: (1) In the neuronal sample as a whole, long segments are more effective than short and. among the short, loud segments are more effective than faint: (2) Units showing their best excitatory response to AS in a certain frequency band are most probably excited by segments with a high proportion of their power concentrated upon or near this band: (3) Units with a slow (build-up) AS response react to a lower number of song segments than those dynamically following AS transients.

Our data give no hint towards adaptive, feature detection properties of single neurons in field L. Instead, these neurons appear to base their response solely on the short-time spectrotemporal structure of the stimulus, irrespective of its natural or artificial origin.

Aves: European starling; Auditory forebrain; Feature detection: Species-specific sound stimuli; Reverse correlation

Introduction

Until the sixties, 'auditory physiology consisted largely of information about the response of the anesthetized brain to simple stimuli such as clicks or pulses of pure tones (Worden and Galambos, 1972). Then, initiated by first reports of feature detector neurons in the visual system (Lettvin et al., 1959), many researchers lost the faith in that these stimuli, and the recording from anaesthetized animals, would allow to disclose the ability of the auditory system under natural conditions. A trend to complex, mainly natural stimuli was supported by studies showing morphological and functional adaptation of the auditory periphery to the spectrum of the species-specific calls in anurans (Frishkopf et al., 1968) and insects (Johnstone et al., 1970). Two major interdisciplinary conferences dealt solely with the recognition of complex acoustic signals and experimental approaches to its investigation (Worden and Galambos, 1972; Bullock, 1977).

Experimental work was mainly done in the cortex of mammals (bat: Suga, 1978; guinea pig: Creutzfeldt et al., 1980; monkey: Newman and Wollberg, 1973; Winter and Funkenstein, 1973; Glass and Wollberg, 1979; Symmes et al., 1980), and in the auditory forebrain of birds (Leppelsack and Vogt, 1976; Scheich, 1977; Scheich et al.. 1979; Leppelsack, 1983; Scheich et al.. 1983; Margoliash, 1983; Müller and Leppelsack, 1985). Two concepts stood in the foreground of such studies: Selectivity and predictability of a single unit's response to natural stimuli. A unit was called selective when it responded only to a small fraction of the complex sound elements tested. In an even less strict way, a unit was termed predictable if the response to a set of natural stimuli can be understood, in its gross structure, from the tuning curve established with pure-tone bursts. In his critical review, Symmes (1981) emphasized the even logical drawbacks of the way these concepts were applied, for instance, that 'selectivity' may describe the stimulus set as much as the responses. Only few papers report efforts to confine the response releasing structures or focal properties

Correspondence to: Gerd J. Dörrscheidt. Lehrstuhl fur Allgemeine Zoologie und Neurobiologie, Ruhr-Universität Bochum. D-4630 Bochum 1. F.R.G.

(Scheich, 1977) in the test sound, either by modifying the natural stimulus or by applying parametric sound models (Suga, 1978; Scheich et al., 1977, 1979; Leppelsack. 1980, 1983; Eggermont and Epping, 1986). Thus, in most cases the basis of the reported selectivity or unpredictability remains unclear.

Predictability in the narrow sense, that is predictability of the time course of the spike discharge, was investigated only by one research group (Eggermont et al., 1983a,b). For midbrain neurons in the frog, they found that the PSTHs predicted for natural stimulation from a wideband-noise responses were at best in qualitative agreement with those actually recorded. Only in part, however, the differences could be attributed to pecularities of the species-specific stimulus, because also the wideband-noise response used for functional characterization was far from perfect correlation with that 'predicted' (Table I in Eggermont et al., 1983a).

In the course of investigations of auditory processing in the thalamus and forebrain of the European starling (Rübsamen and Dörrscheidt, 1986; Bigalke-Kunz et al., 1987: Knipschild et al., 1992, this issue) which were based on a variety of artificial complex test sounds, our conviction increased that the data published on single-unit responses to natural stimuli were not really sufficient to disprove a simple fixed-filter hypothesis. This hypothesis assumes that the spike activity of a single unit (in the restricted behavioural context of electrophysiological experiments) reflects the acoustic stimulus in the very recent past, say the last 100 ms, with a spectrotemporal weight which is independent of the nature of the stimulus, 'artificial' or 'natural'. The model does not exclude that the weight may, as a result of phylo- and ontogenetic influences, favour the filtering of relevant natural sounds over that of laboratory signals.

Our conviction is mainly based on the following elements. First, vocalizations (the most often used natural stimulus) contain manifold acoustic structures as spike releaser candidates. Therefore, it is not surprising to find "unpredictable' neurons which respond to vocalizations and not to simple laboratory stimuli like pure tones. The opposite case, a unit which does not respond to a natural call although it is expected from a careful comparison of the call spectrum and the unit's tuning curve, may also be the consequence of the temporal characteristic already observable in the pure-tone response. Unfortunately, studies including both spectral and temporal aspects in the comparison of responses to artificial and natural stimuli are extremely rare (Symmes et al., 1980; Eggermont et al., 1983a.b). Furthermore, we stated a remaining ambiguity of the findings, in view of the nearly exclusive abstinence from using objective, statistically based methods in stimulus-response analysis.

In the study reported in the present paper, we have tested the fixed-filter hypothesis in single-unit recordings from Field L in the auditory forebrain of the European starling. A complete sequence of the territorial song of a male starling (Schulz et al., 1986) was used as a natural stimulus. This song contains a great variety of acoustic structures (frequency sweeps, noisy and harmonic complexes), and obviously belongs to the acoustic biotope (Aertsen etal., 1979) of the species. Our analyses aim at differences in the response characteristic of single units under natural and artificial stimulation. Primarily, this comparison is based upon the following criteria: (1) The best excitatory frequency range, (2) response delay, and (3) response-intensity dependence. The comparison does not include details of the spike patterns evoked by artificial and natural stimuli. Prediction in the narrow sense (see above) is thus beyond the scope of our approach.

Methods

The study was performed in six adult European starlings *(Sturnus vulgaris),* two males and four females. The procedures followed in surgery, electrophysiological recording from unanaesthetized animals, computer-based on-line stimulus generation, and in the histological verification of the recording sites have been described in Part I of these companion papers (Knipschild et al., 1992 this issue) and by Riibsamen and Dörrscheidt (1986). Also already introduced in part I are two of the new methods (ERD, REX) applied here in stimulus-response analysis. Therefore, the following description can mainly concentrate on the methods required by the inclusion of species-specific calls into our stimulus repertoire.

Artificial stimuli

Artificial test sounds had a duration of 256 ms, including on/off ramps of 12.5 ms each, and were spectrally restricted to preset frequency bands (see Methods of companion paper and Table I). From the rich repertoire used in part I only the following stimulus classes were used with decreasing priority:

BPN = bandpass filtered Gaussian noise (bands 1-10)
SINE = pure tone at the centre-frequency (band 1-9)
RFM = random-walk frequency-modulated carrier (band 1-10)
SWUP = exponential frequency sweep upwards between band limits (band 1-10)
SWDN = the same, but downwards

Activity records made under iso-intensity RFM stimulation, preferably in the best excitatory band of a unit (see below), allow the evaluation of the Effective Response Delay (ERD) which reflects stimulus-response delay widely independent of stimulus nature

(Knipschild et al., 1992 companion paper in this issue). If the ERD is known, sweep-tone responses enable an elegant method for determining a Dynamic Tuning Curve and, by this, the characteristic frequency of a unit (see Data Analysis, and Fig. 5 for an analogue application of the ERD with natural stimuli).

Natural stimuli

As a natural stimulus (NS1) we used a complete sequence of the territorial song of an adult male starling lasting 23.9 s. The vocalization, originally stored on analog tape (Racal Store4), was digitized by the same equipment as used for its later reproduction (Rockland low-pass filter cut-off frequency 9kHz, roll-off 96dB/octave); 12bit AD/DA converter Data Translation DT2801A in a XT-compatible computer, 20000 samples/s). By digital filtering, the resultant data set was then restricted to the frequency range covered by the artificial stimulus set (160-8590 Hz). By removing low-frequency noise and high-frequency alias components, this procedure improved the quality of the replayed song, at least for the human listener.

The acoustical structure of the stored sound was displayed by a computer generated sonagram (Amplitude and Frequency strips in Fig. 2). The display is based on 256-sample (12.8 ms) FFT analyses with Hanning window, calculated every 6.4 ms, giving a frequency resolution of about 64 Hz in the 0 to 10 kHz range. Guided by this display and the acoustical impression during replay, the song was decomposed into 42 segments (Fig. 2). The segments provided the basis for performing quantitative stimulus-response analysis with a set of natural elements of behavioural relevance. For studying the possible influence of song syntax, a second 'natural' stimulus (NS2) was generated by recombining the segments in a shuffled order.

TABLE I

DISTRIBUTION OF BEST BANDS DETERMINED FROM BPN RESPONSES (W= 133 UNITS)

Best Band in Band No	Band Limits [Hz]	No of Units
1	160- 470	8
2	470- 940	10
3	940-1560	20
4	1560-2340	25
5	2340-3 280	24
6	3 280-4380	12
7	4380-5 630	11
8	5 630-7030	6
9	7030-8590	2
10	160-8590	6
No excitatory best response		9

In the experiments, NS1 and NS2 were presented on two intensity levels (75 dB SPL and 55 dB SPL peak intensity) with each stimulus repeated ten times on both levels.

Data analysis

With artificial stimuli, data analysis used many of the methods already applied in the preceding paper:
- dot displays (Fig. 7a,b) for raw-data documentation (On-line and off-line)
- peristimulus-time histograms PSTH (off-line, mainly for latency evaluation)
- ERD (Effective Response Delay).

The best (excitatory) band of a unit was determined with BPN stimuli. It was taken from a response map (Fig. 7c) based on the PAX1 index described in the following. The estimation was routinely cross-checked by visual inspection of the dot display.

In the response model of PAX1 (Pattern index) we assume that the stimulus either affects the number of spikes found in a response window (starting with the stimulus and outlasting it by a defined time span), or that it leads to a non-uniform distribution within this window, expressed by a relative accumulation or reduction of spikes during the stimulus. The flow-chart of Fig. 1 explains the defining quantities and describes the process of selecting one out of nine possible response decisions (including the 'no response' decision marked by a minus sign). Essentially, the procedure consists of the statistical assessment of a two-bin post-stimulus time histogram (PSTH): the first bin is chosen to be congruent with the stimulus period, the second covers the selected post-stimulus period. Possible synchronizing effects of the stimulus within a bin are disregarded which allows, at least in theory, even a single-run response decision.

A statistical decision as to whether a stimulus has influenced the spike count (n_w) in the response window (duration W) requires the null distribution of n_w. It is estimated, for the given number m of stimulus runs, from a sample of spontaneous activity (procedure see Fig. 1). Fractiles of the null distribution (we use the 2%, 10%, 90% and 98% fractiles) serve as thresholds determining the existence and the quality of a response on base of n_w. If n_w does not exceed N_2 or is greater than or equal to N_{98} the algorithm immediately decides for strong inhibition («;) or excitation (»), respectively. Otherwise, the decision is postponed until after the uniformity test.

The second response type considered by PAX1, the non-proportional distribution of spikes between the stimulus period, S, and the post-stimulus section of the window, can be tested without referring to the spontaneous activity. Under the null hypothesis, a spike observed within W will fall into the stimulus period with the probability p = S/W. The decision threshold can

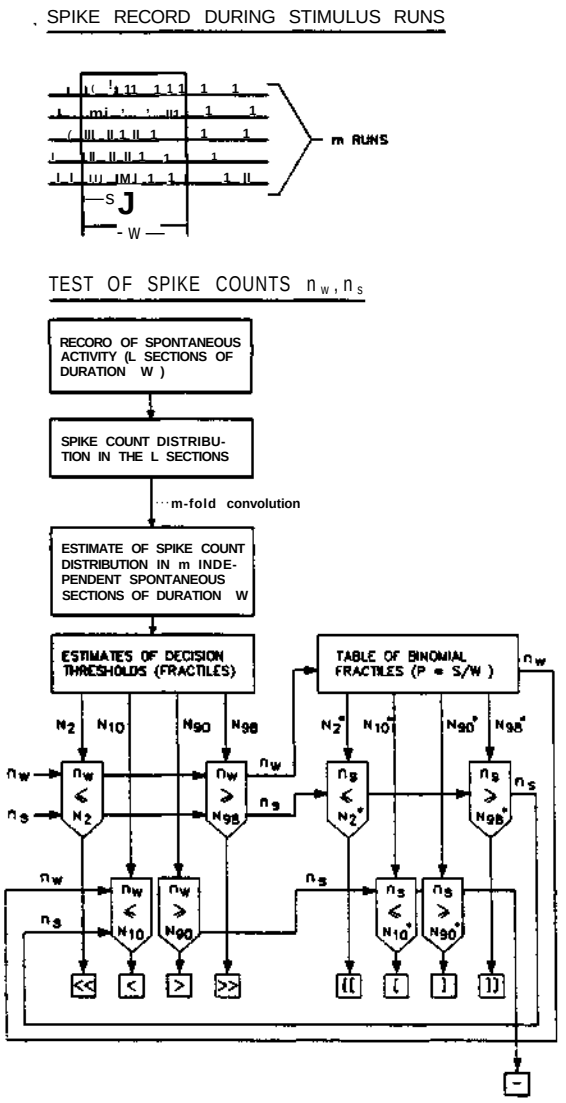

Fig. 1. Flow-chart describing the evaluation of the PAX1 response index. Top: Acquisition of spike counts n_s and n_w during m stimulus runs. S = stimulus duration; W = response window, W > S. Main part: Estimation of decision thresholds $N_{:}$-N_{98} from a record of spontaneous activity. Necessary only once for stable cell recordings and constant (W,S,m). After each parameter test (m runs). n_w is checked against these thresholds and. except the cases $n_w \wedge N_2$ and $^{n}w \wedge \wedge 98$ which indicate strong suppression («;) and strong excitation (») within the response window respectively, n_w serves as a reference for n_s in testing for spike accumulation (],]]) or depletion ([,[[) during the stimulus. Fig. 7 exemplifies the application in the on-line evaluation of response maps.

thus be taken from a computer-stored table of fractiles of the binomial distribution for a sample size n_w. As the flow-chart shows, the process decides for the first response type (<§:,») even if the second test has an equally significant outcome.

In addition to the best band, whenever the necessary data were available a Dynamic Tuning Curve (DTC) was evaluated from the response to exponential frequency sweeps spanning the full range (band 10). For this purpose, onset and stop time of the sweep response was estimated from a narrow-bin PSTH summing-up five runs. With these times corrected by the unit's ERD, the corresponding frequencies were then taken from the frequency-time course of the sweep signal. The low-frequency value of the DTC at the test intensity was calculated as the average of SWUP onset frequency and SWDN stop frequency. With SWUP and SWDN interchanged, the high-frequency boarder of the DTC was found in the same way. Finally, the characteristic frequency (or best frequency) was evaluated by graphical interpolation of the near-threshold DTC values.

Response analysis of the natural stimuli required additional methods. Response strength in segment-related analyses was scaled with the REX index (Knipschild et al., this volume). REX weighs the similarity of spike activity during identical stimulus repetitions. It does not refer to spontaneous activity, and shows only marginal dependence on the spike rate. The analysis period, which is the only parameter which must be specified for REX evaluation, was set congruent to that of the segment.

For identification of response-efficient parameters in the natural song a method was applied based on the backward-correlation principle (deBoer and Kuyper 1968). Specifically, the analysis compares the sound signal in a period preceding the spikes (PESE = pre-event stimulus ensemble; Johannesma, 1972; Aertsen et al., 1980) with the ensemble of sound signals preceding random reference points (SE = stimulus ensemble). Both ensembles are characterized by power spectra: SE by the mean spectrum of the entire stimulus, PESE by a number of short-time spectra centred at different instants t before the reference spike.

In our implementation, analysis was speeded up considerably by using a PSTH representation (binwidth 6.4 ms) of spike activity instead of single spike times. This required to calculate the series of short-time spectra of NS1 (or NS2) only once. As for the sonagrafic display (Fig. 2) we determined the spectra for sections of 12.8 ms (256 sound samples) with hanning weight, overlapping by half, i.e. by 6.4 ms. Spectrum and PSTH were time-related so that the i^{th} spectrum reflects the sound in the i^{th} and the $(i\text{-}1)^{th}$ bin.

Average short-time PESE spectra were calculated for shifts of 0..15 bins, corresponding to 0.96 ms pre-spike time. Following Aertsen et al. (1980) we term the PESE spectra normalized to the SE spectrum the Spectro-Temporal Receptive Field STRF of the unit under the given stimulation. Actually, our STRF plots (Fig. 4) show the dB-difference between these spectra. A positive value at a point (f, -t) in the STRF is

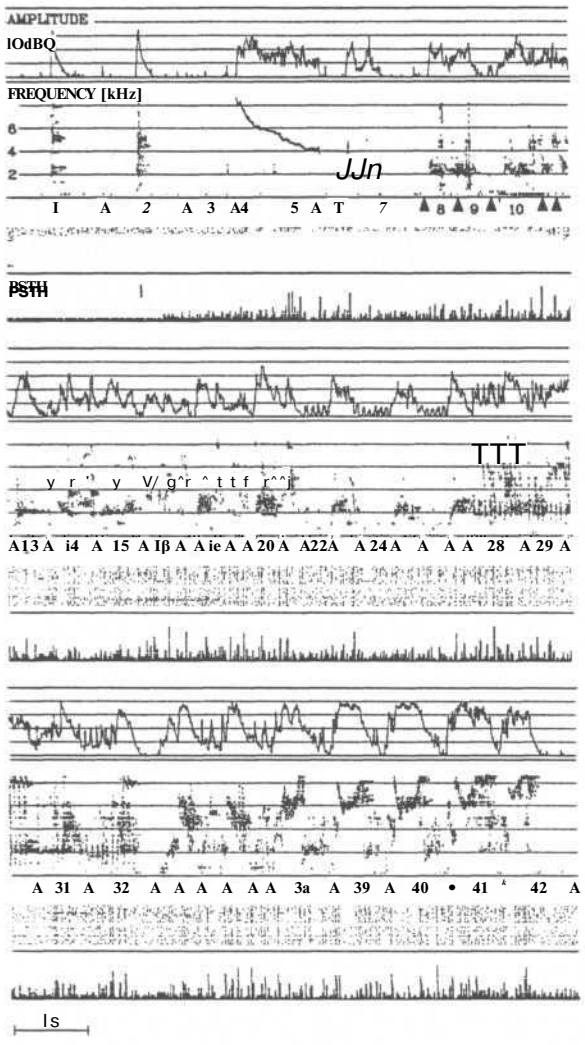

Fig. 2. Off-line generated plot documenting the sound structure of the natural stimulus (NSl) and the spike activity of a single Field L unit recorded during its identical repetition. Total analysis period 23.9 s. Sonagram-like sound description is based on 256-sample (= 12.S ms) FFT analyses with Hanning window, calculated every 6.4 ms. Amplitude course (top): rel. scale, 10 dB/div. Frequency: range 0 to 10 kHz, resolution: 128 components (about 64 Hz), dynamics 0 to - 24 dB relative to the peak in the instantaneous spectrum. Filled triangles below the spike raster mark the borders of the segments used in segment-related analyses (see Methods, Figs. 9 and 10). Spike Raster: response of a single unit (here SF2106) to 10 runs of NSl at 75 dB SPL peak intensity (upper traces) and at 55 dB SPL (below j. PSTH: peri-stimulus-time histogram summing up the spike activity in 12.8 ms bins, full scale: 150 spikes/s in one run.

interpreted as signalling a spike-activating effect of the f-component a period t later. Correspondingly, a negative deflection is interpreted as a sign of suppression *.

* One should be aware that this does not imply a causal relation: in contrast to white-noise stimuli, the short-time spectra of natural stimuli show correlation between the components both within the same and between neighbours. So. an apparently activating component may be only the steady concomitant of the actual spike-efficient component. See Discussion and Aertsen et al. (1981).

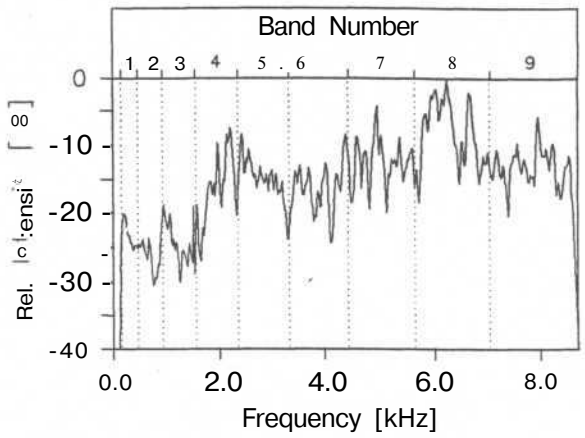

Fig. 3. Mean power spectrum of the natural stimulus sequence NSl and its segment-shuffled version NS2, plotted, relative to its peak value, on a linear frequency scale (bottom). The dotted vertical lines mark the limits of the nine frequency bands to which the spectral energy of the artificial stimuli is restricted (compare Table I). This spectrum serves as reference (SE spectrum) in STRF analyses.

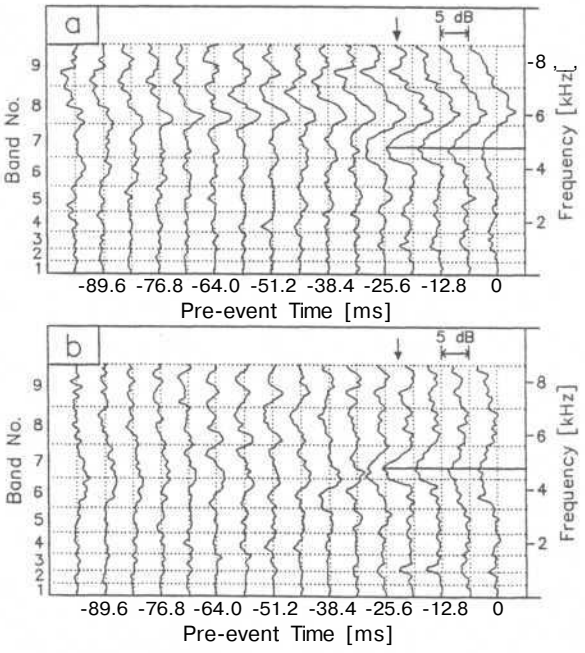

Fig. 4. STRF (spectro-temporal receptive field) of unit SF2106 under stimulation with NSl at (a) 75 and (b) 55 dB SPL peak intensity. The series of 16, vertically oriented short-time spectra describes the period of about 100 ms preceding the spikes in the cell response. In each spectrum, a course near to the vertical, dotted line could be interpreted as a lack of influence of that pre-spike period on spike elicitation. Positive spectral amplitudes (left from the baseline) show an excitatory effect, negative values suppression. For details of the analysis see under Methods. Here, the strongest excitation centres around 5 kHz (right ordinate) or band No.7 (left ordinate). The arrow and the marker line indicate the evaluation of NS response latency (here 21.4 ms) and NS best frequency, respectively, which are compared to analogue parameters of the AS response characteristic (see Figs. 5 and 8). Note that the suppression standing out near 6 kHz in (a) is much more sensitive to the intensity decrease than the excitatory peak.

Results

From a total of 222 neurons recorded, 133 could be subjected to a detailed comparison of their responsiveness to natural and synthetic stimuli. According to histological verification, all analysed units stem, in approximately equal number, from the three layers of field L in the auditory forebrain (Rübsamen and Dörrscheidt 1986). All units, searched with a pure-tone (SINE) sequence, responded to all tested classes of artificial stimuli as well as to both natural stimuli (NS1 and NS2). Referring to the response pattern elicited by the bandpass noise stimulus (BPN) in the individual best frequency band, the majority (66%) of the units were classified as phasic or sustained responders. About half of these units were termed 'transient followers' because their spikes were visibly triggered by the microstructure of the BPN stimulus (e.g. Fig. 7, stimulus band 7). In 21 percent activity increased delayed and slowly (build-up units). The remaining 13% were inhibited by the stimulus, partly followed by rebound excitation after stimulus offset (2%). Spontaneous activity in the sample ranged from 0.5 pps to 122 pps (mean 9.4 pps).

Processing of stimulus frequency

For all 133 units, we tried to determine the best responded frequency range from the responses to bandpass noise (BPN) stimuli, statistically assessed by PAX1 and summarized in response maps (see Methods and Fig. 7). For 124 units, a 'best band' leading to strongest excitation could be specified. The remaining 9 units were inhibited in a wide frequency range. The distribution of the best bands (Table I) obviously mirrors the audiogram of the starling (compare Fig. 1 in part I): more than half of the units have their best band in the frequency range with the lowest behavioural thresholds. Note that 6 units showed the best response to wide-band noise (band 10) rather than to one of the narrow-band stimuli. In most cases, we found the best band fitting into the three-dimensional tonotopic organization of field L (Rübsamen and Dörrscheidt, 1986).

In case of natural song stimulation, we inferred the best excitatory frequency from STRF analysis as the frequency under the maximum positive STRF peak. In Fig. 4 this procedure is shown for unit SF2106. the song response of which was already documented in Fig. 2. (Also some of the following figures will refer to this •model unit'.) Comparison of these best NS frequencies with the AS best bands for all 124 units revealed a nearly perfect congruence. From the 118 units of Table I best responding to one of the narrow-band stimuli (band 1 to 9), 97% (115 units) had their best NS frequency in the same band. The remaining three differed by only one band. The six units of Table I

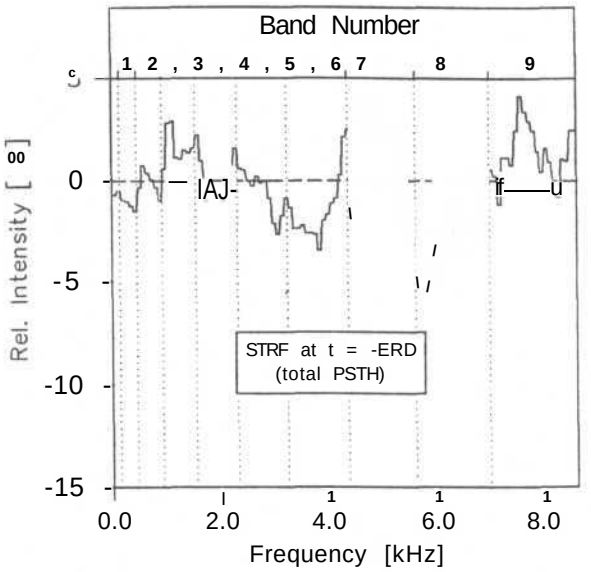

Fig. 5. Cross-section through the STRF for the same unit as in Fig. 4, 75 dB SPL peak intensity. The PESE-spectra of ail spikes recorded were calculated and averaged for a 12.8 ms interval centred at pre-event time -21.4 ms, corresponding to the Effective Response Delay (ERD) evaluated for this unit with RFM stimuli (55 dB SPL). Corresponding to the time axis of Fig. 4, the acoustic delay of 4 ms is not subtracted from the ERD. Shown is the dB-difference of the resulting mean PESE spectrum and the reference (SE spectrum). The diagram discloses a best frequency of 4.6 kHz (band no.7) surrounded by inhibitory sidebands.

which were most strongly activated by the wide-band noise stimulus (band 10) correspondingly showed broad positive plateaus in the STRF analysis.

Obviously, the accuracy of this comparison is restricted by the preset band limits of the artificial stimuli (see Table I) and the also preset time resolution of the STRF analysis (6.4 ms). The latter deficiency could be overcome post experimentum (by choosing a sliding time window for the STRF analysis) while the first needs additional data (and thus experimental time). In 22 units of our sample we were able to record both the response to RFM and to frequency sweeps (SWUP, SWDN) in the unit's best band. This enabled the evaluation of the Effective Response Delay (ERD) (Knipschild et al., this issue), and to pinpoint the best frequency under both stimulus conditions (see Methods for the analysis of sweep responses). In case of the natural stimuli, improved estimates of the best frequency were obtained by determining a cross-section of the spectrotemporal receptive field (STRF) centred at t = - ERD. Fig. 5 shows such analysis for the model neuron. Compared to the course marked in Fig. 4a. the STRF peak appears smaller and more pronounced.

Fig. 6 summarizes the results obtained for all 22 neurons. It reveals a surprisingly good correspondence of the best frequencies under the two stimulus conditions, tonal sweep and natural call. The finding corroborates the foregoing statement based on BPN stimuli.

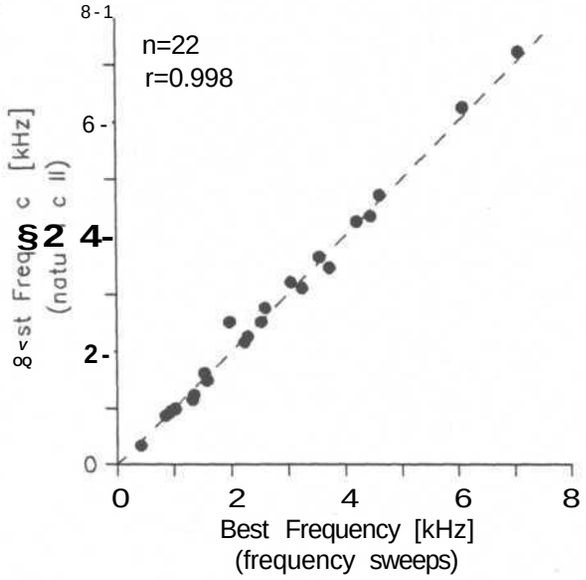

Fig. 6. Comparison of the characteristic frequency evaluated from responses to frequency sweeps (abscissa) and the best spike-eliciting frequency in NS1 responses (ordinate). Data from 22 units. Refer to Methods for the calculation of the characteristic frequency from sweeps. NS best frequency is taken from analyses like Fig. 5. The dashed line marks the 1:1 relation.

For this subset of transient follower neurons, at least, there is no evidence that the great differences in spectrotemporal structure between the two stimulus classes - fast frequency sweeps (1 octave per 45 ms) on one hand, simultaneous frequency complexes on the other - affect the best frequency.

The similarity of the neuronal characteristics inferred from responses to NS and synthetic stimuli is, however, not restricted to the unit's best excitatory frequencies but, in general, encompasses the entire frequency response profile. This shall again be exemplified on unit SF2106. For this unit, careful inspection of the spike raster and the PSTH (Fig. 2) suggests increased spike elicitation by spectral components in the range of 4-5.5 kHz. The STRF analysis of the record supports this interpretation (Fig. 4, Fig. 5) and, additionally, detects inhibitory sidebands centred at 3.5 and 6 kHz. The low-frequent sideband is, however, significant only in the ERD related analysis (Fig. 5). The same frequency dependence is seen in BPN records (Fig. 7a,b), and was also found following sweep up, sweep down and RFM stimulation. Here, best excitation in the range 4.4-5.6 kHz [band No.7] is enclosed by inhibitory sidebands between 2.3-3.3 kHz [band No.5] and 5.6-7.0 kHz [band No.8].

In a further variant of the STRF analysis, we confined the analysis to 'the smallest fragments of the [natural] stimulus that drive the cell as vigorously alone as does the intact call' (Symmes 1981). For this purpose, in each unit the analysis was restricted to the ten sound segments eliciting the strongest excitation, as

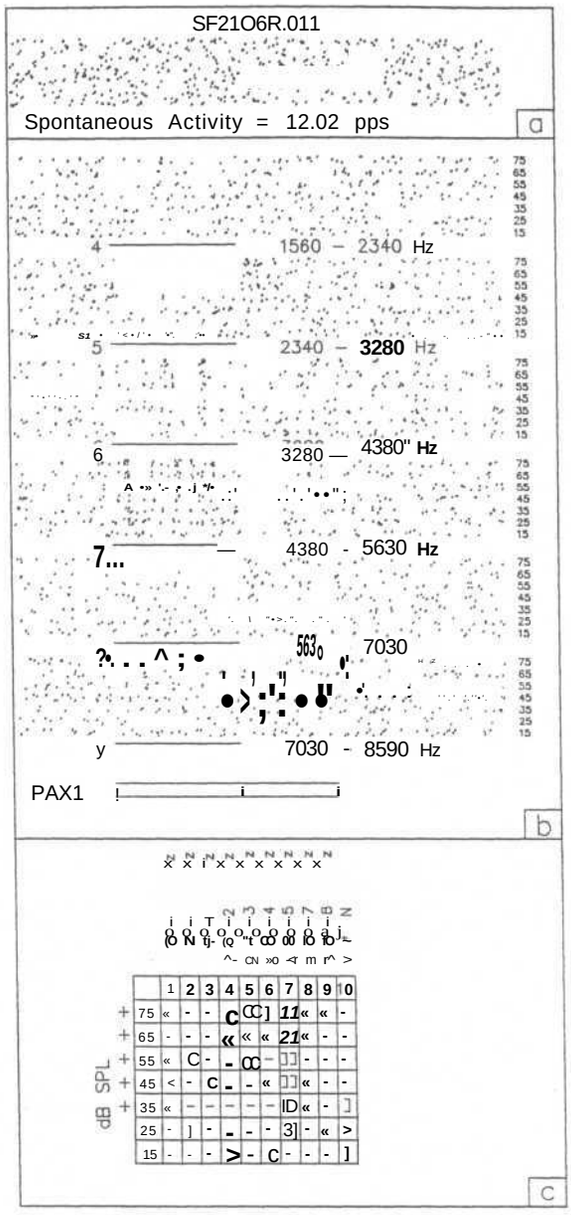

Fig. 7. Dot display of the spontaneous activity of unit SF2106 (a) and its response to bandpass noise (BPN) pulses (b). Modified after the presentation which the experimenter gets on-line on the computer monitor and which is also used as raw-data hardcopy documentation. Response to a BPN stimulus, repeated 5 times in each of the intensities (dB SPL) given at right, is shown on top of the stimulus bar (256 ms) which is enclosed by the band number and the corresponding frequency limits (see Table I). The low-frequency bands below the main response area are omitted, as well as the not efficient band 10 (wide-band). According to the dot pattern the unit was assigned band number 7 as its best band. Note the inhibitory sidebands in band 5 and 8. (c) Response map statistically assessing the spike activity with the PAX1 response index (see Methods). PAX1 gives a rough classification of the five runs performed for each frequency band/intensity pair: <: inhibition, [: inhibition followed by off-excitation, >: excitation,]: excitation followed by off-inhibition,-: no response. Double symbols indicate higher significance *(P < 2%)* than single (10%).

indicated by PSTH peaks (compare Fig. 2). On this basis, ERD related short-time spectra were calculated (Fig. 8) as done for the whole record (Fig. 5). As in the model neuron, also in the cell sample analysed in this way (53 units) the restriction to the most efficient sound segments lets the efficient frequencies stand out in the spectrum. In the sample, STRF peaks differ from the background level by 5 to 50 dB (mean 25.1 dB) which is much larger than the 2 to 19 dB difference (mean 7 dB) found in total record analyses.

Response latency under pure-tone and natural stimulation

The above comparison of frequency selectivity gives no hint towards significant differences between the processing of simple artificial stimuli (BPN, tonal sweeps) and of a complex natural call. This does not exclude, however, that the concurrent frequency components in a natural call could affect a gross temporal parameter, response delay, to which we have access by our methods. Therefore, we compared response latency to pure tones with that to the efficient structures in the natural call. Pure-tone latency was taken from the PSTH of the SINE response in the unit's best band. For the natural stimulus, the pre-event time of the maximum positive STRF peak was used as an adequate measure.

Data for a paired comparison of the two latencies were available for 57 units. No reasonable tone latency could be assigned to nine of them which showed an 'off or 'rebound' response to the 256 ms long stimulus. The same units showed weak STRF maxima near the left margin of the STRF diagram corresponding to NS latencies of about 100 ms. The comparison for the remaining 48 units is visualized in Fig. 9. It reveals that, in all but three units (6%), latencies under the two conditions differ by only a small amount which is in the order of the time resolution used for their evaluation.

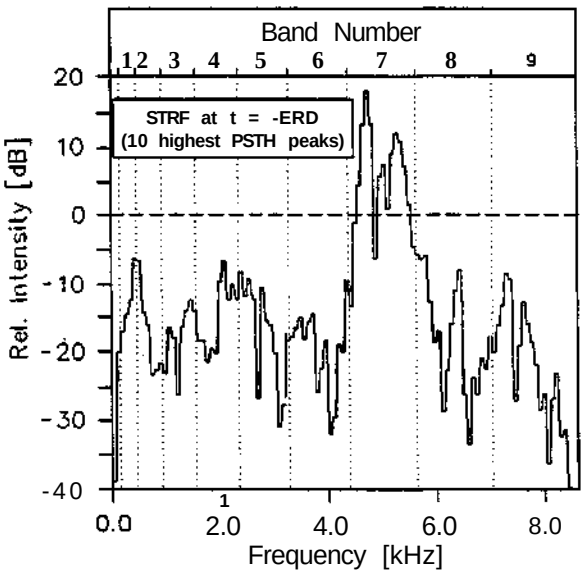

Fig. 8. ERD-corrected STRF analysis of the response of unit SF2106 to the natural call NS1 at 75 dB SPL peak intensity. In contrast to the whole-record analysis (Fig. 5), PESE spectra were only calculated from sound sections directly (+ /- 6.4 ms) neighbouring the ten highest peaks in a narrow-bin (1 ms) PSTH. evaluated from the spike record. Note the difference in the ordinate scale, and that, outside the excitatory frequency range, frequency components are 20 dB less intense than expected without stimulus effect.

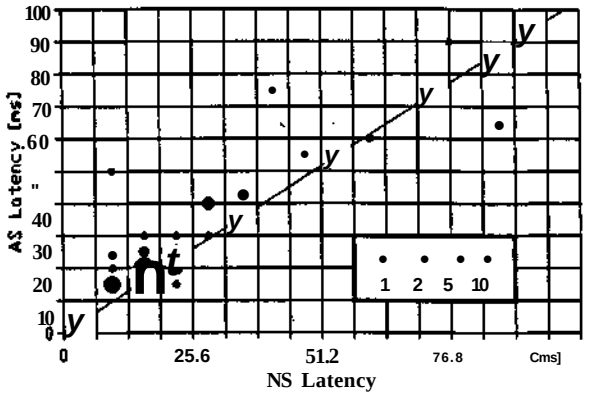

Fig. 9. Response latency under pure-tone and natural call stimulation. Pure-tone latency is taken from the PSTH describing the response at the mid-frequency of the best band (25 runs, average over 35-75 dB SPL) as the lower limit of the first bin (binwidth 5 ms) in which the PSTH exceeds the half-maximal amplitude. An estimate for the latency efficient under natural (NS1) stimulation is drawn from the STRF diagrams (Fig. 4) as the pre-event time of the maximum excitatory deflection plus 3.2 ms (half bin). The latter corrects a bias introduced by the time relation set for the PSTH and the computer sonagram with STRF evaluation. $N = 48$ units (excluding off and rebound responders). The number of units is coded by the area of the filled circles.

Processing of stimulus intensity

With artificial stimuli, intensity dependence was investigated by decreasing intensity from 75 dB SPL, in 10 dB steps, down to threshold (Fig. 7). Response thresholds spanned 15 to 35 dB SPL. Above threshold, in the great majority of the units the stimulus-driven spike rate increased monotonicly with intensity. A direct intra-unit comparison of response-intensity dependence is hampered both by the wide span of sound intensity in the natural call itself (Fig. 2) and by the long recording times required. Due to the latter, we had to restrict the recording of NS responses to two peak intensity levels (55 and 75 dB SPL).

Nevertheless, we could compare 14 units which not only responded over a large intensity range to BPN but had also been tested at both peak intensities with NS1. Response strength was scaled by the stimulus-driven spike rate and by the maximum STRF amplitude, respectively. Twelve units showed monotonic intensity functions with BPN stimuli and a corresponding increase of the STRF peak when NS peak intensity was

raised from 55 to 75 dB SPL. In the remaining two units, exhibiting upper thresholds at 55 and 65 dB SPL respectively, STRF peaks also decreased with NS intensity.

The idea that positive intensity coding is characteristic in Field L also for natural stimuli gains support by an analysis extended over those units in our sample which were tested with NS1 on both peak intensity levels. For each of the 62 units and each of the 42 song segments (Fig. 2) we determined the pair of REX indices (see Methods and part I) attributed to the two intensities. The series of 62 REX pairs obtained for a given segment was then compared in a non-parametric test (Wilcoxon matched pairs signed rank test, $P <$ 2.5%: Siegel, 1956). In 39 segments response strength proved to be significantly higher at a song peak amplitude of 75 dB SPL than at 55 dB. In the remaining three, the trend went into the same direction. Since the segments differ in their sound level by up to 40 dB, the data give evidence, that monotonic intensity coding is typical for the entire suprathreshold intensity range (up to 75 dB SPL) and not only in the range of the tested song peak intensities (55.75 dB SPL).

This test indicates that an 'average neuron' in the sample responds to an increase of intensity by an increase of response strength. Additional analysis has shown that most individual neurons scale the 42 segments on the two sound levels in a consistent way. Among the 62 units, not one showed a qualitative change in it's individual ranking of the segments (rank correlation of the REX indices: positive correlation in all units, in 55 units significant with $P < 1\%$).

Factors determining response efficiency of natural call segments

Segment power and duration

The foregoing analysis of intensity processing lets expect that, in a cell sample, the loudest segments would produce the strongest response. Fig. 10 shows that this expectation is not true without restriction. Relating the segmental REX values to segment power and duration it discloses that the strongest responses are elicited by segments having high power and long duration, the weakest by the low-power (here also short) segments. Essentially, the members in the groups of the most and the least efficient elements remain the same when the song intensity is changed.

Two effects present themselves for explaining the dependence from segment duration. First, unlike pure tones or other unstructered stimuli, the chance for a complex natural stimulus of carrying a trigger for a given neuron will increase with its duration. Second, there is a bias in the definition of REX (see Methods in Knipschild et al., this volume) which, with slowly spiking neurons, favours longer segments.

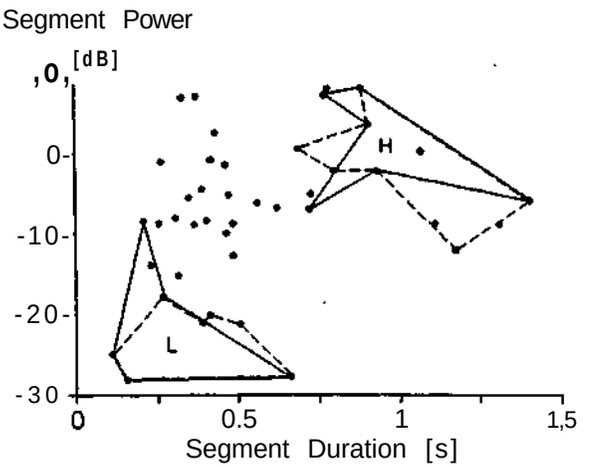

Fig. 10. The influence of segment duration and segment power on the response efficiency of a segment. Each dot represents one of the 42 segments of the natural call (Fig. 2) according to its duration and its power (mean square signal amplitude, 0 dB = mean power of all segments). The segments leading to the strongest responses (H = highest median REX values) and to the weakest (L) responses in a sample of 62 units are connected by solid (test at 75 dB SPL peak intensity) and dashed lines (55 dB SPL).

Segment spectrum

Fig. 6 (and the statistical data given in the text) have revealed concordance between the most spike-efficient frequency in the natural call on one hand, and best bands or best frequencies observed under BPN and sweep stimulation on the other. We therefore studied the relation between the spectral content of a sound segment and the responsiveness of a neuron of a given best band. For four representative song segments, such analysis is documented in Fig. 11. Taken alone, responsiveness possesses strikingly different distributions for the four segments. Comparison with the right column of the figure reveals that these differences reflect the course of the segmental spectrum. A unit will most likely respond if a high proportion of the sound energy concentrates on or near its best band.

Integration time

If latency, as far as it exceeds fibre delay, reflects neuronal integration time, then build-up units should only respond to those segments containing near-bestband power over a sufficient time. This property would restrict the proportion of segments to which such units respond, and so they would appear as more 'selective' than, on the contrary, 'transient followers'.

This in mind, we analyzed NS records of 98 units for the number of responded segments (Fig. 12). As predicted, build-up units were found activated by a significantly smaller number than transient followers (medians = 8/14: $P < 0.05$). So, the higher 'selectivity' of build-up units appears simply as a consequence of their higher integration time observable under artificial stimulation.

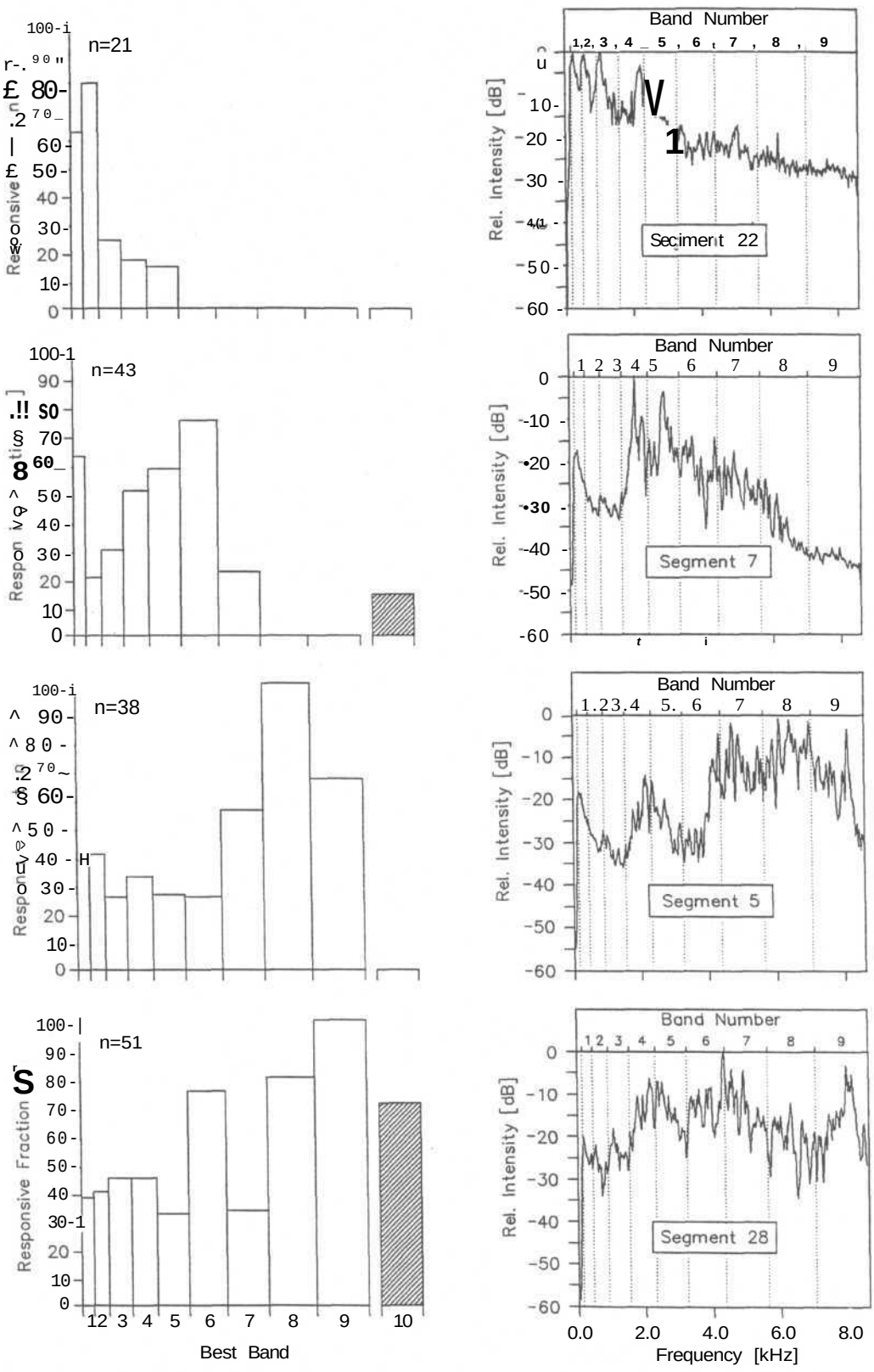

Fig. 11. The relation between the spectrum of a song element and a unit's best band strongly influences the responsiveness of the unit. Concerning their spectral content, the selected segments serve as typical for: (a) a 'low frequency' segment (no. 22), (b) a 'mid frequency' segment (no. 7). (c) a 'high frequency' segment (no. 5) and (d) a 'broadband' segment (no. 28). The histograms on the left show the proportion of units of a given best band (determined in BPN experiments) which respond to this segment. For example: 80% of all neurons having their best band in band 2 (470-940 Hz) respond to segment 22. *N* = total number of units responding to this segment.

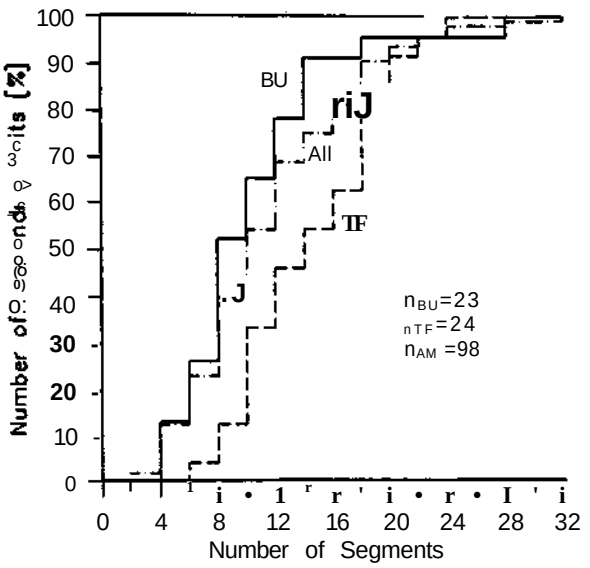

Fig. 12. Relation between the response pattern under BPN stimulation and the number of responded song segments. BU = build-up, TF = transient followers, ALL = whole sample including BU and TF (98 units). The cumulative distribution shows that BU units respond to a significantly less number of segments than TF units.

Recent stimulus history

The natural stimulus in our experiments was an unmodified starling call carrying meaning for a behaving conspecific. Thus it was interesting to look for signs of syntax-dependent responses in the forebrain. As a first step, we compared the response to the natural call (NS1) with that to a stimulus containing the segments (Fig. 2) in randomly permuted order (NS2, see Methods).

Spike records of 28 units were subjected to following analyses:

(a) Responses to NS1 and NS2 as a whole were rated on the basis of STRF analysis. The comparison of the resulting spectra revealed no difference, except in one unit which showed a small latency shift from 38 ms for NS1 to 32 ms in NS2.

(b) In each of the 42 pairs of corresponding segments, spike activity, rated by REX, was compared in a non-parametric test (Wilcoxon matched pairs signed rank test). In a small number of segments (six at 75 dB SPL peak intensity, 4 at 55 dB) the response to the permuted sequence differed significantly *(P<5%)* from that to the original. According to critical inspection of the most differing spike records, however, also these discrepancies could be attributed to response properties already visible in the response to artificial stimuli. Primarily the response to short, poorly efficient segments (see Fig. 10) is influenced if, for instance, the segment preceding in NS2 is responded by strong long-lasting inhibition (in accordance to its BPN response) while the predecessor in NS1 is excitatory or not effective.

(c) For every unit, the similarity in the response to the segments in the two orders was assessed by Spearman rank correlation. In all but one unit, correlation was found highly significant (^< 0.5%), with average coefficients of 0.71 and 0.69 at 75 dB and 55 dB peak intensity, respectively.

Discussion

In this study we have tried to disclose differences between the neuronal response characteristic in the auditory forebrain of the starling, established under stimulation by a species-specific song, and that obtained with more simply structured artificial stimuli. We based this comparison on criteria which refer to both spectral and temporal aspects. Nevertheless, our results give no indication of a specific processing of the natural stimulus. Instead they support the fixed-filter hypothesis stated in the introduction. Spectral (best frequency) and temporal (response latency) parameters of individual neurons were found remarkably similar in both stimulus situations (Figs. 6 and 9). Response efficiency of song elements turned out to be a function of simple parameters (segment power and duration, Fig. 10) and of the match between the spectral composition of the elements and the spectral sensitivity of the neurons (Fig. 11). Furthermore, integration time constants apparent in the response to artificial stimuli allow at least a rough prediction whether a unit will belong to the 'selective' responders of the natural stimulus sequence (Fig. 12).

Acoustic sensitivity of field L neurons

We have found that, within the limits of field L explored in multi-unit recordings (Riibsamen and Dörrscheidt, 1986), all units are auditory and respond both to a wide range of artificial stimuli (see Methods) and to at least some segments of the conspecific territorial song (Fig. 12). This seems in contrast to Leppelsack and Vogt (1976) who have found, in also unanaesthetized preparation, a large number of non-auditory units interspersed into the auditory area which responded neither to artificial stimuli nor to a sequence of starling vocalizations. Reasons for this discrepancy are difficult to pinpoint. Most likely it will contribute to the difference that we scanned the entire hearing range of the species in frequency and intensity, and that we applied response tests sensitive also to time-locking (REX, PAX1).

Processing of stimulus frequency

Our data have shown a nearly perfect concordance between the best excitatory frequency band under artificial and natural stimulation (Fig. 6 and text). Furthermore, we have reported a general agreement of the

frequency profiles in both cases. The latter requires some comments.

Being aware of the problems connected with the interpretation of spectro-temporal receptive fields when established with other stimuli than white noise (see Methods), we performed cross and auto-correlation analyses of the natural call (NS1), based on the power courses in the nine frequency bands and on the time shifts used in STRF analysis. The analyses have revealed that correlation is non-negative and extends only over neighbouring bands. The non-white call spectrum will thus possibly pretend additional spike-releasing frequencies, especially near the main frequency component. It may also complicate the detection of spike-inhibiting frequencies, but it seems legal to accept clear negative deflections in the STRF (as in Fig. 4a) as actual inhibitory influences. We do now understand why we found the best correspondence between the STRF and the dot display of BPN responses for units (mainly from the input region L2) which showed inhibitory sidebands like the model unit (Figs. 3 and 6).

Such correspondence supports the idea of a fixed spectro-temporal filter weighing the sound stimulus irrespective of its context. Further support comes from the finding (Fig. 11) that the responsiveness of a unit for a given sound segment follows essentially the sound level concentrated in its best band. The best bands being tonotopically arranged in field L (Guinea fowl: Bonke et al., 1979; Starling: Müller and Leppelsack, 1985; Rübsamen and Dörrscheidt, 1986), this also means that the spectrum of a sound segment is mapped into a spatial distribution of response probability. It fits with this model that, in field L of the Guinea fowl, the highest score of 'Iambus' responding units is found in regions where the dominant spectral band of this call (1-2.3 kHz) coincides with the pure-tone map (Scheich et al.. 1979). Such mapping can also be seen in 2DG autoradiography (Scheich et al., 1983).

Response latency

Fig. 9 shows a surprising correspondence between the response latencies evaluated for pure-tone and natural stimuli (NS1). It is even more striking because of the different methods used. Within the, admittedly poor, time resolution reflecting the bin-widths used in PSTH and STRF analysis we see a one-to-one increase of AS and NS latency. The remaining absolute difference of about 3 ms by which, on average, NS latency is shorter, is probably due to the difference in the adaptive state of the involved neuronal net in both tests. This effect was already made responsible for differences between ERD and pure-tone latency (Knipschild et al.. this volume). There, in a cell sample for which both latency indices could be evaluated, the ERD was also found shorter than the pure-tone latency at the same intensity level, with a median difference of 6 ms.

Response versus intensity

According to our analyses, intensity processing in field L neurons is the same for artificial (mainly BPN) and natural stimuli. In the tested range (up to 75 dB SPL) monotonic characteristics prevail. For the natural call this was shown by comparing the response strength evoked by its 42 segments at two intensity levels, differing by 20 dB (p. 239). When a direct comparison, NS versus BPN response, was possible the parallelism of AS and NS intensity processing showed up also for units having upper thresholds.

Which properties make up response efficiency of a song segment?

Given, as a rule, a monotonic intensity characteristics, we can expect, for each field L neuron and every segment, that an increase in intensity will increase response strength. (This expectation appears especially legitimate when the REX index is used for response scaling because it primarily weighs spike synchronization by the gross and fine structure of the stimulus while being insensitive against spike rate saturation.) However, response strength itself is largely predetermined by the spectro-temporal structure of the stimulus and its fit to the response characteristics of the individual neuron. This is expressed by the individual ranking which every neuron establishes for the segmental sequence and which remains largely unaffected by a 20 dB intensity change (p. 239). As a consequence for the whole cell sample, the group of the most effective segments remains stable, as does that of the least effective (Fig. 10). The spectral sensitivity of a neuron, roughly characterized by its best band, largely determines which segments out of the 42 can elicite a response. The result of the analyses comparing segment spectra and best-band specific responsiveness (see Fig. 11 for representative graphs) suggests as a working hypothesis that the proportion of the segment power concentrating on or near the best band of a unit decides upon the response.

One should be surprised, however, that we can establish such relation between the responsiveness of a unit of a given best band and the overall power spectrum of a sound segment, because the spectrum cannot reflect the short-time structure of the sound unequivocally. (Compare the spectrum of segment 5 in Fig. 11 suggesting high-frequency bandpass noise, and its sonagraphic description in Fig. 2 disclosing a nearly pure frequency sweep.) We can, for instance, not infer from the power spectrum whether, in a short-time analysis, frequency components inside the best band stay long enough for evoking a response even in build-up units. In fact, such units respond to significantly less segments than their counterpart, the transient followers (Fig. 12). Given nearly equal distributions of the best bands for both functional types, this difference indi-

cates the limits of a purely spectral interpretation of responsiveness. On the other hand, this result shows that the inclusion of temporal parameters, deducible from responses to simple stimuli, takes a further bit of secret from 'selective' units. Our experiments comparing, in the same unit, the response to the original starling call with that to a segment-shuffled version have revealed no or only minor differences. This holds both for the assessment of the entire spike record by the STRF, and for the comparison on a segmental base using REX. As stated, REX weighs mainly similarity of the spike discharge in repeated stimulus runs. Therefore, we cannot exclude that the form of the segment-related PSTH may differ under the two conditions whereas REX does not. PSTHs of those units and segments, which were scaled with the greatest REX differences and which were therefore subjected to a more detailed analysis, differed, however, only in their leading parts. We expect this restriction even more when REX differences are small.

Conclusion

Our results provide good reasons to answer the provocative question asked in the subtitle to the negative. The stimulus range releasing a response in field L neurons has proven wide enough to allow their isolation and to explore their general response characteristic by a fairly small repertoire of artificial stimuli. The spectral sensitivity disclosed by bandpass noise stimuli or frequency sweeps is in good agreement with that when processing a complex conspecific song. This agreement allows to predict to what song section the neuron will respond and, to some degree, response strength. Units showing their main sensitivity in a frequency range represented with low power in the natural stimulus will respond to a smaller number of elements than those with spectral sensitivity more in the centre of the song. The same relation seems to hold in the time domain among slowly integrating units and transient followers. When selectivity of an auditory neuron is defined on the proportion of responded elements in a potpourri it will thus find a simple explanation.

However, as already stated in the Introduction, our comparison of the response characteristics does not include details of the spike pattern. We only dealt with the existence of a response to a sound element and with its relative strength. Therefore, the analysis is not able to show whether NS responses can be predicted in the narrow sense. Concerning field L, we are hopeful that the latter task could be tackled successfully too. This, however, would require a different approach in the selection of the artificial stimulus set which, in the present paper, aimed mainly at the basic parameters describing frequency selectivity and response latency.

Which relevance have the results for feature detection concepts?

The wide stimulus range releasing a response in the cell sample excludes that the sample contains candidates for 'pontifical' alias 'grandmother' neurons which, in feature detection concepts (Bullock, 1961, 1986; Worden and Galambos, 1972; Scheich, 1977, 1983; Suga, 1978; Ewert, 1987), mark one extreme, that of highly specialized or 'multiple tuned' units. We did, however, not really expect to find experimental evidence for their existence (see Symmes, 1981). The functional role of the units described here appears as that of moderately tuned filter neurons imposing to complex stimuli individual spectro-temporal weights. Thereby, the spectral rating, expressed in the best frequency range of a unit, shows a clear map in field L (Rübsamen und Dörrscheidt, 1986). As maps in general, the tonotopic map facilitates the imagination that sound recognition could be performed by a sensory-motor interface (Scheich, 1983) which derives motor commands from the scanned output of the filter neurons.

Acknowledgement

We are indebted to Petra Knipschild for preparation of the graphs. This work was supported by the Deutsche Forschungsgemeinschaft (SFB 114).

References

Aertsen, A.M.H.J., Smolders, J.W.T. and 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.

Aertsen, A.M.H.J., Johannesma, P.I.M. and Hermes, D.J. (1980) Spectrotemporal receptive fields of auditory neurons in the grass-frog. II. Analysis of stimulus-event relation for tonal stimuli. Biol.Cybern. 38, 235-248.

Aertsen, A.M.H.J., Olders, J.H.J, and Johannesma, P.I.M. (1981) Spectro-temporal receptive fields of auditory neurons in the grassfrog. Biol. Cybern. 39, 195-209.

Bigalke-Kunz, B., Rübsamen, R. and Dörrscheidt, G.J. (1987) Tonotopic organization and functional characterization of the auditory thalamus in a songbird, the European starling. J. Comp. Physiol A 161, 255-265.

Bonke, D., Scheich, H. and Langner, G. (1979) Responsiveness of units in the auditory neostriatum of the guinea fowl *(Numida meleagris)* to species-specific calls and synthetic stimuli. I. Tonotopy and functional zones of field L.J. Comp. Physiol. 132, 257-276.

Bullock, T.H. (1961) The problem of recognition in an analyzer made of neurons. In: W.A. Rosenblith (Ed.), Sensory Communication. MIT Press, Cambridge, MA, pp. 717-724.

Bullock, T.H. Ed. (1977) Recognition of complex acoustic signals. Life Sciences Res. Rep. Vol.5. Dahlem Konferenzen, Berlin.

Creutzfeldt, O., Hellweg, F.C. and Schreiner, Chr. (1980) Thalamocortical transformation of responses to complex auditory stimuli. Exp. Brain Res. 39, 87-104.

DeBoer, E. and Kuyper, P. (1968) Triggered correlation. IEEE Transact. Biomed. Engin. BME 15, 169-179.

Eggermont, J.J., Aertsen, A.M.H.J. and Johannesma, P.I.M. (1983a) Quantitative characterisation procedure for auditory neurons based on the spectro-temporal receptive field. Hear. Res. 10, 167-190.

Eggermont, J.J., Aertsen, A.M.H.J. and Johannesma, P.I.M. (1983b) Prediction of responses of auditory neurons in the midbrain of the grass frog based on the spectro-temporal receptive field. Hear. Res. 10, 191-202.

Eggermont, J.J. and Epping, W.J.M. (1986) Sensitivity of neurons in the auditory midbrain of the grassfrog to temporal characteristics of sound. III. Stimulation with natural and synthetic mating calls. Hear. Res. 24, 255-268.

Ewert, J.P. (1987) Neuroethology of releasing mechanisms: Prey-catching in toads. Behav. Brain Sci. 10, 337-405.

Frishkopf, L.S., Capranica, R.R. and Goldstein, M.H., Jr. (1968) Neural coding in the bullfrog's auditory system - A teleological approach. Proc. IEEE 56, 969-980.

Glass, I. and Wollberg, Z. (1979) Lability in the response of cells in the auditory cortex of squirrel monkeys to species-specific vocalizations. Exp. Brain Res. 34, 489-498.

Johannesma, P.I.M. (1972) The pre-response stimulus ensemble of neurons in the cochlear nucleus. In: B.L. Cardozo (ed.X Proc. IPO Symp. on hearing theory. Eindhoven, pp. 58-69.

Johnstone, B.M., Saunders, J.C and Johnstone, J.A. (1970) Tympanic membrane response in the cricket. Nature 227, 625-626.

Knipschild, M., Dörrscheidt, G.J. and Riibsamen, R. (1992) Setting complex tasks to single units in the avian auditory forebrain. I. Processing of complex artificial stimuli. Hear. Res. 57, 216-230.

Leppelsack, H.J. (1980) Response selectivity of auditory forebrain neurons in a songbird. In: R. Nöhring (Ed.) Acta XVII Congr. Intern. Ornithol. 1978, Dtsch. Ornithol. Ges., Berlin, pp. 728-734.

Leppelsack, H.J. (1983) Analysis of song in the auditory pathway of song birds. In: J.P. Ewert, R.R. Capranica and D.J. Ingle (Ed.), Advances in Vertebrate Neuroethology. Nato ASI series (Life Sciences Vol. 56), Plenum, NewYork, pp. 783-799.

Leppelsack, H.J. and Vogt, M. (1976) Responses of auditory neurons in the forebrain of a songbird to stimulation with species-specific sounds. J. Comp. Physiol. 107, 263-274.

Lettvin, J.Y., Maturana, H.R. McCulloch, W.S. and Pitts, W.H. (1959) What the frog's eye tells the frog's brain. Proc. IRE 47, 1940-1951.

Margoliash, D. (1983) Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. J. Neurosci. 3, 1039-1057.

Miiller, CM. and Leppelsack, H.J. (1985) Feature extraction and tonotopic organization in the avian auditory forebrain. Exp. Brain Res. 59, 587-599.

Newman, J.D. and Wollberg, Z. (1973) Responses of single neurons in the auditory cortex of squirrel monkeys to variants of a single call type. J. Exp. Neurol. 40, 821-824.

Riibsamen, R. and Dörrscheidt, G.J. (1986) Tonotopic organization of the auditory forebrain in a songbird, the European starling. J. Comp. Physiol. A 158, 639-646.

Scheich, H. (1977) Central processing of complex sounds and feature analysis. In: T.H. Bullock (Ed.), Recognition of complex acoustic signals. Life Sciences Res. Rep. Vol.5. Dahlem Konferenzen, Berlin, pp. 161-182.

Scheich, H., Bock, W., Bonke, D., Langner, G. and Maier, V. (1983) Acoustic communication in the guinea fowl *(Numida meleagris)*. In: J.P. Ewert, R.R. Capranica and D.J. Ingle (Ed.), Advances in Vertebrate Neuroethology. Nato ASI series (Life Sciences Vol.56), Plenum, New York, pp. 731-782.

Scheich, H., Langner, G. and Bonke, D. (1979) Responsiveness of units in the auditory neostriatum of the guinea fowl *{Numida meleagris)* to species-specific calls and synthetic stimuli. II. Discrimination of iambus-like calls. J. Comp. Physiol. 132, 257-276.

Scheich, H., Langner, G. and Koch, R. (1977) Coding of narrow-band and wide-band vocalizations in the auditory midbrain nucleus (MLD) of the Guinea fowl *(Numida meleagris)*. J. comp. Physiol. 117, 245-265.

Schulz, W., Rübsamen, R., Dörrscheidt, G.J. and Abs, M. (1986) Analyse von Gesängen individuell erkannter Stare. J. Ornithol. 127, 380.

Siegel, S. (1956) Nonparametric Statistics for the Behavioral Sciences. McGraw-Hill, New York.

Suga, N. (1978) Spezialization of the auditory system for reception and processing of species-specific sounds. Fed. Proc. 37, 2342-2354.

Symmes, D. (1981) On the use of natural stimuli in neurophysiological studies of audition. Hear. Res. 4, 203-214 (1981)

Symmes, D., Alexander, G.E. and J.D. Newman (1980) Neural processing of vocalizations and artificial stimuli in the medial geniculate body of squirrel monkeys. Hear. Res. 3, 133-146.

Winter, P. and 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, 498-504.

Worden, F.G. and Galambos, R. Eds. (1972) Auditory processing of biologically relevant sounds. Neurosci. Prog. Res. Bull. 10. NRP Brookline, MA.