Syntax processing by auditory cortical neurons in the FM–FM area of the mustached bat *Pteronotus parnellii*

(nonspeech signals/communication calls/auditory cortex/hearing/mammals)

Karl-Heinz Esser*†‡, Curtis J. Condon*§, Nobuo Suga*, and Jagmeet S. Kanwal*¶

*Department of Biology, Washington University, St. Louis, MO 63110; †Department of Comparative Neurobiology, University of Ulm, Ulm, D-89069, Germany; and ¶Institute for Cognitive and Computational Sciences, Department of Neurology, Georgetown University Medical Center, Washington, DC 20007

ABSTRACT Syntax denotes a rule system that allows one to predict the sequencing of communication signals. Despite its significance for both human speech processing and animal acoustic communication, the representation of syntactic structure in the mammalian brain has not been studied electrophysiologically at the single-unit level. In the search for a neuronal correlate for syntax, we used playback of natural and temporally destructured complex species-specific communication calls—so-called composites—while recording extracellularly from neurons in a physiologically well defined area (the FM–FM area) of the mustached bat's auditory cortex. Even though this area is known to be involved in the processing of target distance information for echolocation, we found that units in the FM–FM area were highly responsive to composites. The finding that neuronal responses were strongly affected by manipulation in the time domain of the natural composite structure lends support to the hypothesis that syntax processing in mammals occurs at least at the level of the nonprimary auditory cortex.

From a linguist's point of view (1–3), "syntax" denotes a rule system that accounts for the ability to produce an infinite variety of sequences (i.e., words and sentences) from a fixed number of phonemes (i.e., vowels and consonants). Apart from human speech, rule systems for the sequencing of species-specific vocalizations have been found repeatedly in both birds (4, 5) and nonhuman mammals (6, 7). Thus, more generally, "syntax" can be understood as any system of rules that allows one to predict the sequencing of *communication* signals (3).

Our present report of syntax processing by auditory cortical neurons in the mustached bat *Pteronotus parnellii* is founded on several previous studies in this species by employing both neurophysiological and behavioral methods. First, in the context of intraspecific acoustic communication, mustached bats frequently combine otherwise independently emitted simple syllables to form either isosyllabic trains or heterosyllabic composites. The syntactical rules for the generation of the last-mentioned higher order constructs have been revealed in detail (8). Second, the auditory cortex of *P. parnellii* is arguably the most intensively studied and best understood of all mammals (9). Thus, established representational maps (e.g., ref. 10) can be used as a reference to record from defined areas and functional subtypes of auditory cortical neurons. Third, neurons in the FM–FM area of the mustached bat auditory cortex were shown recently to respond facilitatively to isosyllabic pairs (11). Presumably, these neurons mediate acoustic communication (11) in addition to their primary (i.e., first discovered) function in echolocation (12).

FM–FM neurons exhibit heteroharmonic combination-sensitivity for paired stimuli mimicking the FM components of the bats' echolocation sounds (pulse and echo) (12) and often respond well to communication calls (11). Thus, we hypothesized that FM–FM neurons may also show combination-sensitivity for heterosyllabic composites (8), providing a neuronal correlate for syntax of communication sounds in a mammalian species.

MATERIALS AND METHODS

Details of the surgical and chronic recording techniques have been described elsewhere (13). Briefly, the care and use of animals in these experiments were approved by the Animal Care and Use Committee of Washington University. During the 4 to 8-h recording session, the awake bat was continuously monitored via video for signs of discomfort or distress. Extracellular recordings were obtained from the FM–FM cortical area of mustached bats ($n = 5$) using custom-made lacquer-coated tungsten microelectrodes (tip diameter 4–8 μm). Single units were isolated with a level and time window discriminator (BAK Electronics) on the basis of spike height and slope of the waveform. The composite calls used as stimuli were digitized, manipulated, and played back using an A/D-D/A converter board (Data Translation, DT2821-G-8DI; 250-kHz sample rate; 12-bit amplitude resolution) and SIGNAL software (Engineering Design) (14). Acoustic signals mimicking the FM components of the bats' echolocation sound (pulse) and echo were generated with conventional analog equipment (13).

Upon isolation of a single unit, the FM_1–FM_n subtype (i.e., FM_1–FM_2, FM_1–FM_3, or FM_1–FM_4; Fig. 1) and best pulse–echo delay (BD) of the neuron were determined audiovisually. The neuron's responses to the most effective biosonar pulse–echo pair were further quantified by peristimulus-time histograms (PSTH) using a computer and customized data acquisition software (MI² System). The particular FM_1–FM_n neuron subtype, the unit's BD, and the location of the recording site in relation to superficial landmarks such as characteristic blood vessels on the brain surface were used for comparison with established representational maps (10, 12). We then sequentially presented the 10 digitally stored and energy-matched (on the basis of root-mean-square values) natural composites (Fig. 2) at a rate of 1/s. This composite series was played 10 times at each stimulus level (50, 60, 70, 80, and 90 dB sound pressure level), and the neuron's response to each of the composites was measured as the number of spikes minus spontaneous activity in a 20-ms window around the peak response. The "best composite" was determined by summing the neuron's responses over five stimulus levels. Subsequent testing of the neuron's response selectivities was carried out using the best composite at its best amplitude. Depending on the particular best composite and/or the duration of the neuronal response,
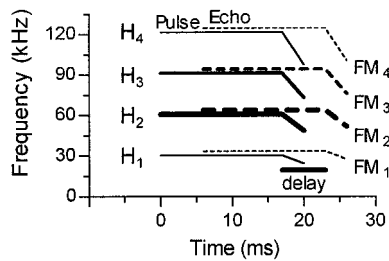
FIG. 1. Schematized sonagram of the mustached bat orientation pulse (solid lines) and the Doppler-shifted, time-delayed echo (dashed lines). The four harmonics (H$_1$–H$_4$) of both pulse and echo each contain a long constant frequency component and a short FM component (FM$_{1-4}$). The thickness of the lines indicates differences in relative amplitude of harmonics. According to their heteroharmonic combination-sensitivity for echolocation pulse–echo pairs, three major types of FM$_1$–FM$_n$ neurons can be distinguished: FM$_1$–FM$_2$, FM$_1$–FM$_3$, and FM$_1$–FM$_4$ (redrawn from ref. 15).

spike numbers (spontaneous activity subtracted) were measured using either a 100- or 200-ms time window. The best composite (AB) and its individual syllables (A and B) were presented separately (200 trials each, repetition rate = 1/s) to determine the neuron's response ratio in percent—i.e., the response to the original composite divided by the sum of the responses to the individual parts and the quotient multiplied by 100. To address both facilitative and suppressive intersyllable interactions, the term "response ratio" was preferred to the more commonly used "facilitation ratio." *Temporal facilitation* was indicated if the response to the entire composite was >120% of the sum of responses to the heterosyllabic parts, whereas *temporal suppression* was defined as a response to the original <80% of the sum of the responses to the individual syllables. To examine the effect of temporal order on the neuron's response selectivity, the original composite was played backwards, reversed in order (BA), and/or a silent period was inserted between the two syllables of the composite. The responses to these manipulations were compared with the neuron's response to the original composite. Also, interactions between spectral components of composites were studied by filtering the syllables. Filtered syllables were presented both separately (A or B) and recombined (AB) to determine the neuron's response ratio. For frequency domain processing the terms *spectral facilitation* and *suppression* were used according to the above definitions for temporal combination-sensitivity.

RESULTS

Preference for Composite Communication Calls. The response selectivity for composite communication calls was investigated in 107 FM–FM neurons with best pulse–echo delays ranging from 2.1 to 15 ms. Different types of FM–FM neurons (FM$_1$–FM$_2$, FM$_1$–FM$_3$, FM$_1$–FM$_4$) responded best to different types of composites (Fig. 3). More than half of the FM$_1$–FM$_2$ neurons (n = 20 of 36, Fig. 3) preferred one of two calls (Fig. 2; (no. 8) sHFM-QCFs, (no. 10) dRFM-cDFM) containing biosonar-like signals (i.e., downward FMs), whereas 67% of the FM$_1$–FM$_3$ neurons responding to composites (n = 18 of 27) preferred one of three calls (Fig. 2; (no. 3) fSFM-QCFs, (no. 5) fSFM-bUFM, (no. 6) fSFM-bUFM-TCFs) dominated by multi-harmonic sinusoidal FMs. In contrast, FM$_1$–FM$_4$ neurons (n = 35) had a more even distribution of the "best composite" (Fig. 3). No response to composite call stimuli was obtained from 14% of the units studied.

Time Domain Processing. A kernel density plot (nonparametric estimator; not shown) of the response ratios of 92 neurons (Fig. 4) revealed a trimodal distribution, indicating that the population of neurons responding to composite communication calls was divisible into three groups that matched the qualitative criteria of suppression, unchanged, and facili-
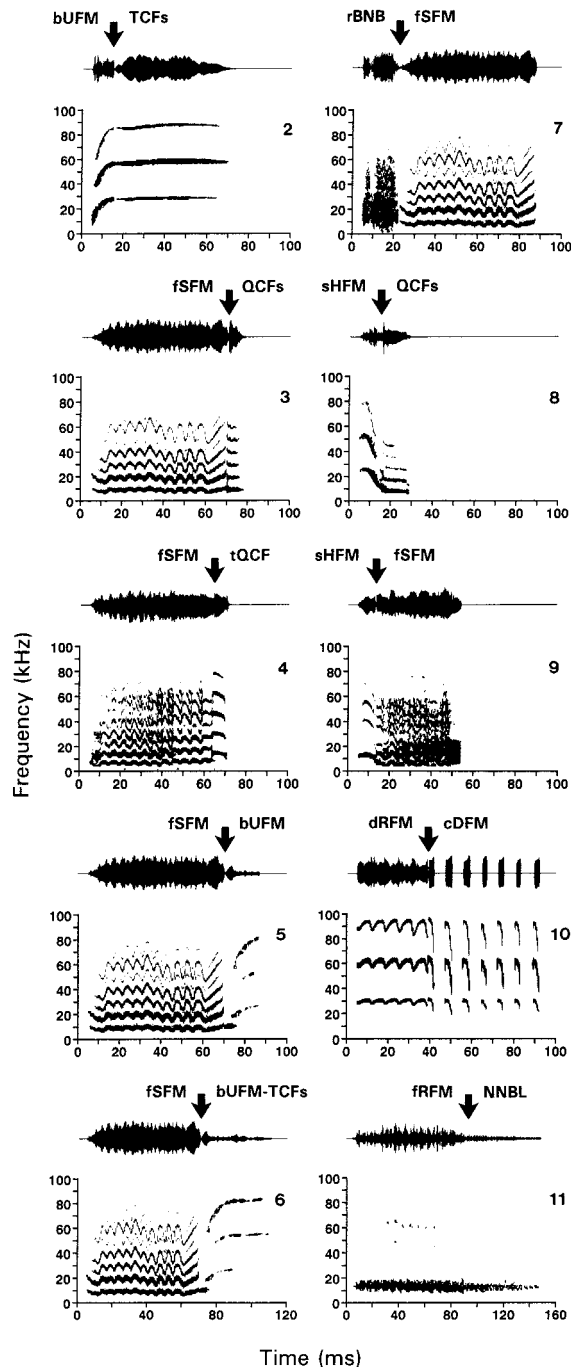

FIG. 2. Oscillograms (*Upper*) and sonagrams (*Lower*) of composite communication calls (nos. 2–11) of the mustached bat *P. parnellii*. Composites are made up of two (all except no. 6) or three (no. 6) distinct components (syllables) that the bats combine without an intervening silent interval (connecting points of syllables indicated by arrows). The composites (e.g., bUFM-TCFs, no. 2) are named according to the simple syllables (e.g., bent Upward FM-short True Constant Frequency) from which they are composed (for complete nomenclature, see ref. 8). For determining a unit's spontaneous activity a no-stimulus interval (i.e., no. 1, not shown) was used at the beginning of each composite series.

tation. Post-priori tests of significance were run by performing a normal fit to the three separate distributions of response ratios bounded by values corresponding to the notches in the trimodal distribution. t tests generated P values of <0.01 between adjacent populations. These boundaries closely match (i.e., 116% vs. 120%) the "arbitrary" criteria conventionally adopted to define facilitation.
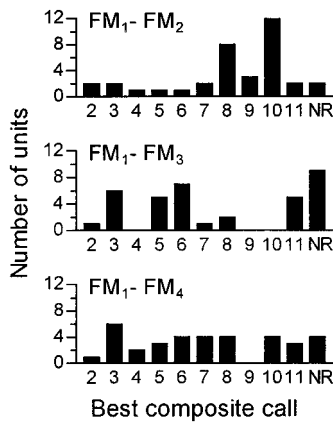
FIG. 3. Neurons of different FM–FM subtypes (FM_1–FM_2, FM_1–FM_3, FM_1–FM_4) showed preferences for different composite calls (nos. 2–11 in Fig. 2). Some units were unresponsive (NR) to composite call stimuli, but could be driven by acoustic stimuli mimicking the particular harmonics of the bat's echolocation pulses and echoes.

As shown in Fig. 4 and summarized in Table 1, 21% of FM–FM neurons showed temporal facilitation in response to their particular best composite (range of response ratios = 122-1108%). For example, the neuron in Fig. 5 (*B–D*) responded vigorously with facilitation (response ratio = 304%) to composite dRFM-cDFM (Fig. 5 *A* and *B*) but responded poorly to the individual syllable components (Fig. 5 *C* and *D*). On the other hand, 30% of all FM–FM neurons tested showed temporal suppression (Fig. 4, Table 1; range = 35–77%), i.e., the response to the entire composite (e.g., Fig. 5 *E* and *F*) was <80% of the sum of the responses to the components (Fig. 5 *G* and *H*). In this example, the initial syllable (sHFM) of composite sHFM-fSFM inhibited the neuron's response to the second syllable (fSFM, Fig. 5 *E* and *F*) by 56%, to which it responded robustly when presented alone (Fig. 5*H*).

The importance of the temporal structure of the composite calls became most evident when playing these stimuli in
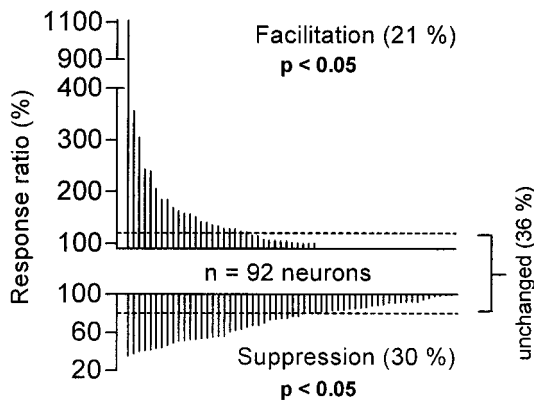

FIG. 4. Distribution and magnitude of temporal facilitation and suppression (for criteria, see *Materials and Methods*) of FM–FM area auditory cortical neurons responding to their particular best composite call. Vertical lines indicate response ratios of individual neurons ($n = 92$). Responses of neurons whose response ratios fell between the corresponding criteria (i.e., 80–120%) were considered as "unchanged." For statistical comparison, the Mann–Whitney U test (two-tailed) was chosen because response ratios of neurons were not distributed in a Gaussian manner. For neurons with response ratios of >100% ($n = 32$, *Upper*) spike numbers (spontaneous activity subtracted) elicited by composite stimuli (i.e., AB) were compared with the sums of responses elicited by the particular simple syllables in isolation (i.e., A+B). A corresponding analysis was performed on spike numbers of neurons with response ratios of <100% ($n = 58$, *Lower*). Within each group of neurons, differences in spike numbers (AB vs. A+B) were found to be statistically significant as indicated in the figure.

Table 1. Summary of the temporal combination sensitivity of FM–FM neuron subtypes (FM_1–FM_n) for composite calls

Response type	FM_1–FM_2	FM_1–FM_3	FM_1–FM_4	$\Sigma\ FM_1$–FM_n
Facilitation, %	8	42	11	21
Suppression, %	31	19	40	30
No change, %	56	14	37	36
No response, %	6	25	11	14
n	36	36	35	107

reverse, introducing a silent period between the syllables, or reversing the order of syllables within a composite. Except for two neurons studied (nos. 7 and 29; $n = 2$ of 21), responses to the reversed composite call were always reduced as compared with responses to the corresponding original composite (Fig. 6). For instance, neuron no. 102 showed temporal facilitation for composite fSFM-bUFM (Fig. 7 *A* and *B*) but almost failed to respond to the individual components when presented separately (Fig. 7 *C* and *D*) or even to the entire composite call if played reversed (Fig. 7*E*). Also, other manipulations in the time domain of composite structure (see above) typically resulted in the loss of a facilitated response or even in the complete loss of a response. Correspondingly, simply introducing a silent period ≥ 0.5 ms between both syllables of the composite fSFM-bUFM (Fig. 7 *A* and *F*) resulted in a progressive decay of neuronal response. Finally, at intersyllable silent intervals ≥ 3 ms, unit no. 102 (Fig. 7*F*) completely ceased
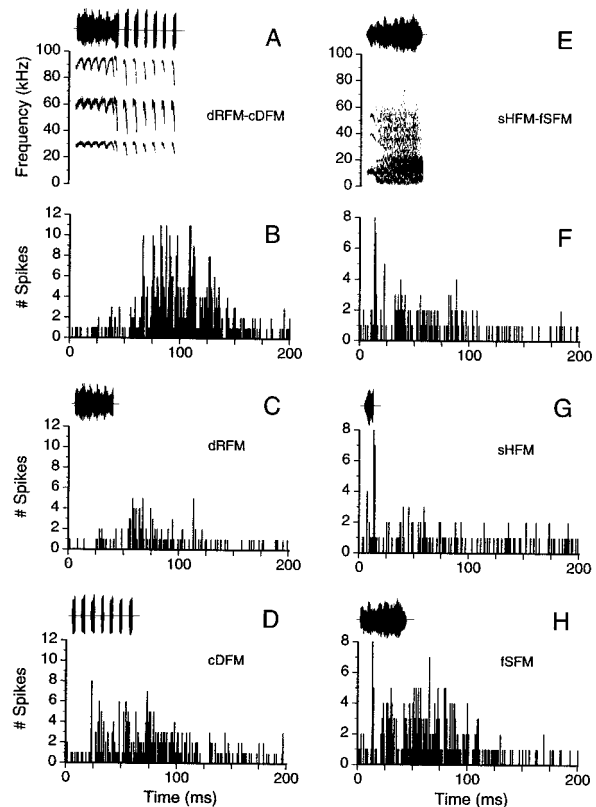

FIG. 5. Left column (*A–D*), temporal facilitation (response ratio = 304%). (*A*) Oscillogram and sonagram of composite no. 10 (dRFM-cDFM), the best composite for unit no. 1 (FM_1–FM_2, BD = 6 ms). (*B*) PSTH (bin width = 1 ms) shows unit's strong response to the original composite (dRFM-cDFM). (*C*) PSTH shows unit's weak response to the first syllable (dRFM). (*D*) PSTH shows unit's weak response to the second syllable (cDFM). Right column (*E–H*), temporal suppression (response ratio = 44%). (*E*) Oscillogram and sonagram of composite no. 9 (sHFM-fSFM), the best composite for unit no. 25 (FM_1–FM_2, BD = 5.3 ms). (*F*) PSTH shows unit's response to original composite (sHFM-fSFM). (*G*) PSTH shows unit's response to the first syllable (sHFM). (*H*) PSTH shows unit's strongest response was to the second syllable (fSFM).
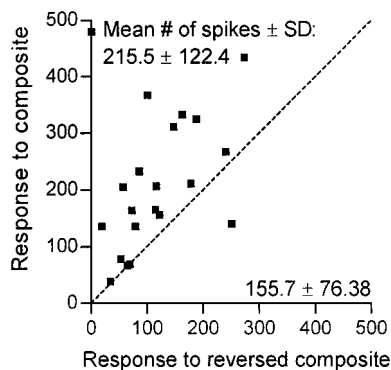
FIG. 6.   Effect of playing composite call stimuli in reverse. For 21 neurons (■; range of response ratios = 35–1108%), responses to the particular best composite and the corresponding stimulus played in reverse could be studied quantitatively. The scatter plot indicates total spike numbers (200 stimulus repetitions less spontaneous activity) for each stimulus condition and neuron. For each neuron, the time window for spike counting was individually adjusted (range = 40–200 ms) to fully include both responses to forward and backward stimulation (dashed bisector = no-change level). Differences in spike numbers (mean ± SD) are statistically significant (two-tailed paired t test, $P < 0.01$), indicating that responses to reversed composites are significantly reduced compared with the original.

to respond. Reversing the order of syllables also resulted in the loss of a facilitated response (compare Fig. 8 *B* vs. *H*).

Frequency Domain Processing. Spectral combination sensitivity of FM–FM neurons for composite stimuli was found to coexist with temporal combination sensitivity. For example, neuron no. 94 (Fig. 8) did not respond well to individual syllables (Fig. 8 *C* and *D*), to the reversed composite (Fig. 8*G*), or when the syllables were presented in the reversed temporal order (Fig. 8*H*). For maximum excitation, this neuron required, in addition to the correct time structure, a specific combination of spectral bands, namely the low frequencies of the first syllable (<35 kHz) and the high frequencies (>65 kHz) of the second (Fig. 8 *E* and *F*), a combination of spectral bands that overlapped with the neuron's FM$_1$–FM$_n$ subtype (i.e., FM$_1$–FM$_3$). Interestingly, in more than half of the neurons studied (56%, n = 23 of 41), we found such facilitative interactions between spectral components of the composite calls. Generally, the excitatory frequency bands that mediated the spectral facilitation were related to the specific FM$_1$–FM$_n$ subtype (n = 2, 3, or 4). Only 15% (n = 6) of the FM–FM neurons studied showed spectral suppression, and the remaining 12 neurons (29%) showed neither spectral facilitation nor suppression.

Biosonar vs. Composite Call Processing. As a rule, highest facilitation ratios of FM–FM neurons were found when using synthetic pulse–echo stimuli presented at the neuron's particular BD. The average facilitation ratio for biosonar signals was 518 ± 330% ($\bar{X}$ ± SD) compared with 228 ± 223% for composite calls (two-tailed paired t test, n = 19 units with temporal facilitation in response to their particular best composite, $P < 0.01$). Further, the temporal pattern of the neuronal response differed for biosonar and communication call stimuli. The average latency of the peak response in the PSTH was significantly longer for the unit's best composite call (79.9 ± 31.9 ms) compared with the average response latency for pulse–echo pairs (24.9 ± 6.1 ms; two-tailed paired t test, n = 19 units with temporal facilitation in response to their particular best composite, $P < 0.0001$). Such long latencies in response to composite calls indicate that FM–FM neurons are integrating composite communication calls over a much longer time period compared with biosonar signals.

DISCUSSION

In echolocating bats, target distance information is encoded in the time delay between the emitted biosonar pulse and the
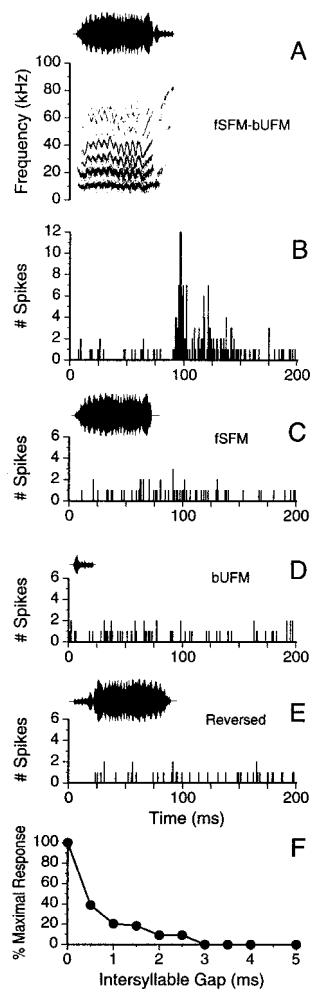

FIG. 7.   (*A*) Oscillogram and sonagram of composite no. 5 (fSFM-bUFM), the best composite for unit no. 102 (FM$_1$–FM$_3$, BD = 8.1 ms). (*B*) PSTH shows unit's robust response to the original composite (fSFM-bUFM; response ratio = 1108%). (*C*) Unit almost failed to respond to the first syllable (fSFM). (*D*) Unit nearly failed to respond to the second syllable (bUFM). (*E*) No response to playback of the reversed composite. (*F*) Unit's response decreased dramatically and eventually ceased when a silent period ≥0.5 ms was inserted between the two components of composite fSFM-bUFM.

returning echo. Correspondingly, neurons in the FM–FM area of the mustached bat auditory cortex are tuned to particular combinations of FM components in orientation sounds and echoes (i.e., FM$_1$–FM$_2$, FM$_1$–FM$_3$, or FM$_1$–FM$_4$) separated by a specific delay (12, 13). Hence, responses of FM–FM neurons (12, 13) are dependent on both spectral composition and temporal order of sound signals.

Microchiropteran bats—e.g., the mustached bat—also produce a variety of communication sounds. From a detailed analysis of communication call structure in *P. parnellii*, the presence of strong constraints on the use of simple syllables as components of composites is obvious (8). Thus, of the 342 disyllabic combinations in composites possible in theory, less than 15 have been found to occur (8). The present observations that neuronal responses to these composite communication calls were highly vulnerable to (*i*) reversal of order of syllables within a natural composite, (*ii*) introducing a silent period between the syllables, and (*iii*) playing the stimulus in reverse provide independent lines of evidence at the single unit level that such syntax in communication sounds is processed by neurons in the mammalian nonprimary auditory cortex. These findings, together with the occurrence of both facilitative and
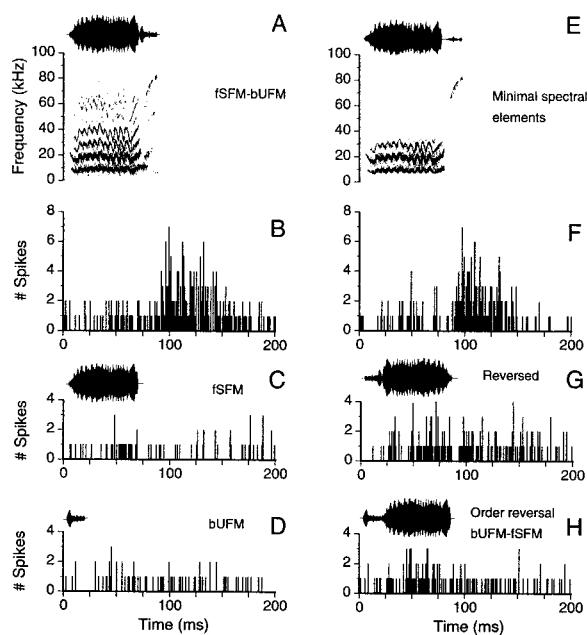
FIG. 8. (*A*) Oscillogram and sonagram of composite no. 5 (fSFM-bUFM), the best composite for unit no. 94 (FM$_1$–FM$_3$, BD = 6.4 ms). (*B*) PSTH shows unit's robust response to the original composite (fSFM-bUFM; response ratio = 242%). (*C*) Unit failed to respond to the first syllable (fSFM). (*D*) Unit failed to respond to the second syllable (bUFM). (*E*) Oscillogram and sonagram of the filtered composite (first syllable, low-pass filtered at 35 kHz; second syllable, high-pass filtered at 65 kHz). (*F*) PSTH shows unit's robust response to minimal spectral elements of composite fSFM-bUFM necessary to excite this unit to a level equivalent to the original. (*G*) PSTH shows unit's weak response to the reversed composite. (*H*) PSTH shows unit's weak response to the reversed order of syllables of composite no. 5 (i.e., BA).

suppressive intersyllable interactions, clearly point to the importance of syntax for processing communication calls.

In the forebrain of a songbird (white-crowned sparrow), even more specific temporal combinations of sound elements were found necessary for maximum excitation of "song-specific" units (16). These rare hyperstriatum ventrale pars caudale (HVc, ref. 17) [or higher vocal center (18)] neurons responded best to the bird's *own* (autogenous) song, whereas other songs, even of the same dialect, elicited weak or essentially no excitation (16, 19). As in our sample of units exhibiting temporal combination sensitivity for heterosyllabic composites, responses of "song-specific" neurons were strongly affected by experimental manipulations of temporal characteristics of the stimulus sequence (16). In the zebra finch HVc, neurons even generally seemed to prefer autogenous as compared with conspecific songs (20). As in songbirds, neurons exhibiting selectivity for the individual's own communication calls might reside in the mustached bat's cortex. However, so far, individually distinct sound characteristics (and neurons tuned to the corresponding acoustic parameters) were described only for this species' echolocation pulses (21). Therefore, further behavioral and neurophysiological studies in the mustached bat are of great importance. Given auditory cortical neurons were found to be highly sensitive to spectrotemporal features emerging *de novo* from the combination of different simple syllables to composite communication calls, the behavioral significance of these signals is a crucial issue for future research. The question arises whether or not we will succeed in developing an experimental design (for comparison see ref. 22) to demonstrate different behavioral responses to playback of composite calls in forward and backward direction or to other complex stimulus manipulations presently described.

The cellular mechanisms underlying neuronal selectivity for temporal order of sound signals have been studied most elegantly by *in vivo* intracellular recordings in the auditory forebrain of the zebra finch (23, 24). Here, burst-firing nonlinearity and long-lasting hyperpolarization were recognized as major mechanisms to integrate auditory context (23, 24). In mammals, the auditory cortex is thought to serve as a substrate for complex temporal processing, including temporally extended processing of brief acoustic signals (12, 25–27). The present finding that ≈50% of FM–FM neurons showed complex intersyllable interactions in the time domain (either facilitation or suppression) lends support to the hypothesis that syntax processing also occurs at the level of the (nonprimary) auditory cortex. However, in contrast to birds, corresponding data from *in vivo* intracellular recordings in the mammalian auditory forebrain are not available (for relevant work on the auditory midbrain, see ref. 28). From both intracellular (23, 24) and extracellular (16, 20) studies in the songbird, it is evident that HVc neurons can integrate auditory information over hundreds of milliseconds. Interestingly, facilitated responses of FM–FM neurons to composite stimuli reveal a similarly increased integration time compared with integration times for signals mimicking the bats' biosonar. Apart from different durations, the fact that composite communication calls are more complex, much less stereotyped, and hence less predictable than echolocation signals might account for this difference. Moreover, in the bird HVc, neuronal responses to song seem hardly affected by the insertion of a short silent period between phrases (16), unlike FM–FM neurons responding facilitatively to composite stimuli (Fig. 7). Such silent intervals between individual syllables (phrases) are characteristic of bird songs in nature (e.g., ref. 29) but absent in composite communication calls of behaving mustached bats (8).

Despite its obvious significance, the neuronal processing of acoustic sequences at the auditory cortical level has not been studied extensively, neither in humans nor in any other mammal except for a few species of bats (for review, see ref. 9). Among the latter, only in the mustached bat (ref. 11 and present study) have stimuli mimicking species-specific vocalizations (i.e., communication calls) other than echolocation pulse–echo pairs been employed systematically. In the cat auditory cortex, neurons were found to respond facilitatively to paired tones if presented at fairly brief (e.g., 300–600 ms) intervals (25, 30). The majority of units studied were affected by tonal contour (i.e., the neurons' responses differed depending on whether ascending, descending, or nonmonotonic tone sequences were presented) and/or serial position of stimuli in multitone sequences (25). With respect to time domain processing, cat vocalizations have not been employed quantitatively as acoustic stimuli in studies of auditory cortex. In the squirrel monkey, the response selectivity of auditory cortical neurons was studied by using playback of natural, reversed, or temporally destructured species-specific vocalizations (e.g., refs. 31 and 32). Although call-responsive units could be classified as "generalists," "specialists," or in between according to the number of vocalizations to which they responded (33, 34), no clear effect of reversing the communication calls was found (31), and it was not possible to single out precisely the acoustic features determining a cell's response (32, 33). Studies in the macaque auditory cortex have provided promising results that may lead to a more detailed understanding of time domain processing of primate communication calls. However, it is still too early to make generalized conclusions from these experiments (35).

In humans, temporal order of perceptual elements according to semantic and syntactical constraints is a prerequisite for the intelligibility of speech (36). Further, acoustic sequencing in music leads to perceptions such as rhythm, melody, and chroma (i.e., interval-size patterning) (25, 37). In the case of human speech processing, intraoperative recording experiments are of necessity too few and limited for thorough characterization of

neuronal filter properties (38). Thus, interspecies comparisons of syntax processing (human/mustached bat) are not possible at the single unit level. However, perceptual features that have previously been thought to be speech-specific, such as categorical perception (39, 40), perceptual constancy despite variability in many acoustic dimensions (41), perception of the formant structure in multi-tone complexes (41), and phoneme perception (42), are gaining acceptance as general preadaptations for the analysis and recognition of communication sounds in mammals, including humans (41). Thus, it seems possible that, similar to the mustached bat auditory cortex (present study), combination- or multi-combination-sensitive neurons are also involved in the perception of some parameters of speech, such as syntax, in humans.

Facilitative interactions observed between *spectral* components (e.g., harmonics) of composites are consistent with previous findings in the species' FM–FM area (for details, see refs. 11, 12, and 14). Typically, the excitatory frequency bands that mediated the spectral facilitation were related to the specific FM_1–FM_n neuron subtype. Thus, it might be argued that instead of performing syntactical analysis, these combination-sensitive cortical neurons might be fortuitously stimulated by combinations of sounds embedded in the natural composite structure. However, as revealed by the experiments where composite communication calls were played in reverse (see Fig. 6), clearly neuronal responses to composites cannot be predicted from the spectral characteristics of the calls alone but are greatly influenced by temporal signal characteristics. As discussed above, the influence of the temporal context on composite communication call processing even extends across syllable boundaries. Similar to the mustached bat (present study), evidence for nonlinear summation of spectral energy by auditory cortical neurons in the primary auditory cortex of the cat (43) and the nonprimary auditory cortex of the macaque (35) is based on studies using playback of species-specific communication calls segmented in the frequency domain (for review of relevant work on other bat species, see ref. 9).

As indicated by different response selectivities of FM–FM neuron subtypes in the mustached bat (Fig. 3), a heterogeneous set of neurons responsive to composite communication calls is embedded in a mapped representation of echolocation pulse–echo delays. This heterogenity is evident only when different types of species-specific communication calls are used as experimental acoustic stimuli. In our paradigm for unraveling composite processing by auditory cortical neurons, we selected FM–FM neurons responding consistently over a broad range of sound intensities. Such a "level tolerance" (44) of neuronal response selectivities (i.e., independence of the distance of the sound source) can be regarded as a prerequisite for auditory cortical units mediating acoustic communication. Our study provides evidence at the single cell level that such neurons are involved in the processing of syntax of communication calls.

1. Chomsky, N. (1957) *Syntactic Structures* (Mouton, The Hague).
2. Umiker-Sebeok, J. & Sebeok, T. A. (1980) in *Speaking of Apes*, eds. Sebeok, T. A. & Umiker-Sebeok, J. (Plenum, New York), pp. 1–59.
3. Snowdon, C. T. (1982) in *Primate Communication*, eds. Snowdon, C. T., Brown, C. H. & Petersen, M. R. (Cambridge Univ. Press, Cambridge, U.K.), pp. 212–238.
4. Balaban, E. (1988) *Behavior* 105, 292–322.
5. Marler, P. & Peters, S. (1988) *Ethology* 77, 125–149.
6. Cleveland, J. & Snowdon, C. T. (1982) *Z. Tierpsychol.* 58, 231–270.
7. Tembrock, G. (1977) *Tierstimmenforschung* (Ziemsen, Wittenberg Lutherstadt, Germany), pp. 28–46.
8. Kanwal, J. S., Matsumura, S., Ohlemiller, K. & Suga, N. (1994) *J. Acoust. Soc. Am.* 96, 1229–1254.
9. O'Neill, W. E. (1995) in *Hearing by Bats*, eds. Popper, A. N. & Fay, R. R. (Springer, New York), pp. 416–480.
10. Suga, N. (1990) *Neural Networks* 3, 3–21.
11. Ohlemiller, K. K., Kanwal, J. S. & Suga, N. (1996) *NeuroReport* 7, 1749–1755.
12. O'Neill, W. E. & Suga, N. (1979) *Science* 203, 69–73.
13. Suga, N., O'Neill, W. E., Kujirai, K. & Manabe, T. (1983) *J. Neurophysiol.* 49, 1573–1626.
14. Ohlemiller, K. K., Kanwal, J. S., Butman, J. A. & Suga, N. (1994) *Auditory Neurosci.* 1, 19–37.
15. Suga, N. (1988) in *Auditory Function*, eds. Edelman, G. M., Gall, W. E. & Cowan, W. M. (Wiley, New York), pp. 679–720.
16. Margoliash, D. (1983) *J. Neurosci.* 3, 1039–1057.
17. Nottebohm, F., Stokes, T. M. & Leonard, C. M. (1976) *J. Comp. Neurol.* 165, 457–486.
18. Nottebohm, F., Alvarez-Buylla, A., Cynx, J., Kirn, J., Ling, C.-Y., Nottebohm, M., Suter, R., Tolles, A. & Williams, H. (1990) *Phil. Trans. R. Soc. Lond. B* 329, 115–124.
19. Margoliash, D. (1986) *J. Neurosci.* 6, 1643–1661.
20. Margoliash, D. & Fortune, E. S. (1992) *J. Neurosci.* 12, 4309–4326.
21. Suga, N., Niwa, H., Taniguchi, I. & Margoliash, D. (1987) *J. Neurophysiol.* 58, 643–654.
22. Esser, K.-H. (1997) in *Schizophrenia and Psychoacoustics*, eds. Nielzén, S. & Olsson, O. (Lund University Press, Lund, Sweden), in press.
23. Lewicki, M. S. & Konishi, M. (1995) *Proc. Natl. Acad. Sci. USA* 92, 5582–5586.
24. Lewicki, M. S. (1996) *J. Neurosci.* 16, 5854–5863.
25. Weinberger, N. M. & McKenna, T. M. (1988) *Music Percept.* 5, 355–390.
26. Ravizza, R. J. & Belmore, S. M. (1978) in *Handbook of Behavioral Neurobiology*, ed. Masterton, R. B. (Plenum, New York), Vol. 1, pp. 459–501.
27. Condon, C. J., Galazyuk, A., White, K. R. & Feng, A. S. (1997) *Auditory Neurosci.* 3, 269–287.
28. Covey, E., Kauer, J. A. & Casseday, J. H. (1996) *J. Neurosci.* 16, 3009–3018.
29. Marler, P. & Peters, S. (1989) in *The Comparative Psychology of Audition*, eds. Dooling, R. J. & Hulse, S. H. (Lawrence Erlbaum Associates, Hillsdale, NJ), pp. 243–273.
30. McKenna, T. M., Weinberger, N. M. & Diamond, D. M. (1989) *Brain Res.* 481, 142–153.
31. Glass, I. & Wollberg, Z. (1983) *Brain Behav. Evol.* 22, 13–21.
32. Wollberg, Z. & Newman, J. D. (1971) *Science* 175, 212–214.
33. Winter, P. & Funkenstein, H. H. (1973) *Exp. Brain Res.* 18, 489–504.
34. Newman, J. D. & Wollberg, Z. (1973) *Brain Res.* 54, 287–304.
35. Rauschecker, J. P., Tian, B. & Hauser, M. (1995) *Science* 268, 111–114.
36. Flanagan, J. L. (1972) *Speech Analysis Synthesis and Perception* (Springer, New York), pp. 276–321.
37. Carterette, E. C. & Kendall, R. A. (1989) in *The Comparative Psychology of Audition*, eds. Dooling, R. J. & Hulse, S. H. (Lawrence Erlbaum Associates, Hillsdale, NJ), pp. 131–172.
38. Creutzfeldt, O., Ojemann, G. & Lettich, E. (1989) *Exp. Brain Res.* 77, 451–475.
39. Ehret, G. & Haack, B. (1981) *Naturwissenschaften* 68, 208–209.
40. May, B., Moody, D. B. & Stebbins, W. C. (1989) *J. Acoust. Soc. Am.* 85, 837–847.
41. Ehret, G. (1992) in *The Auditory Processing of Speech: From Sounds to Words*, ed. Schouten, M. E. H. (Mouton de Gruyter, Berlin), pp. 99–112.
42. Kuhl, P. K. & Padden, D. M. (1982) *Percept. Psychophys.* 32, 542–550.
43. Watanabe, T. & Katsuki, Y. (1974) *Jpn. J. Physiol.* 24, 135–155.
44. Suga, N. (1992) *Phil. Trans. R. Soc. Lond. B* 336, 423–428.