(Orthoptera: Tettigoniidae: *Conocephalus nigrapleurum*). *Can. J. Zool.*, **56,** 1479–1487.
Otte, D. 1977. Communication in Orthoptera. In: *How Animals Communicate* (Ed. T. A. Sebeok), pp. 334–361. Bloomington, Indiana: Indiana University Press.
Otte, D. & Loftus-Hills, J. 1979. Chorusing in *Syrbula* (Orthoptera: Acrididae). Cooperation, interference, competition, or concealment? *Entomol. News*, **90,** 159–165.
Reynolds, W. J. 1980. A re-examination of the characters separating *Chorthippus montanus* and *C. parallelus* (Orthoptera: Acrididae). *J. Nat. Hist.*, **14,** 283–303.
Riede, K. 1983. Influence of the courtship song of the acridid grasshopper *Gomphocerus rufus* L. on the female. *Behav. Ecol. Sociobiol.*, **14,** 21–27.
Skovmand, O. & Pedersen, S. B. 1978. Tooth impact rate in the song of a shorthorned grasshopper: a parameter carrying specific behavioural information. *J. comp. Physiol.*, **124,** 27–36.
Skovmand, O. & Pedersen, S. B. 1983. Song recognition and song pattern in a shorthorned grasshopper. *J. comp. Physiol.*, **153,** 393–401.

(Received 21 April 1986; revised 4 June 1986; MS. number: sc-301)

Combinatorial Animal Communication with Computable Syntax: Chick-a-dee Calling Qualifies as 'Language' by Structural Linguistics

One of the fundamental differences between animal communication and human language is that the latter is combinatorial: a small set of phonemes generates a huge set of morphemes, which in turn are grouped to make higher lexical units. Wilson (1975: page 188) envisioned animal signal-units A, B and C generating the combinations A, B, C, AB, AC, BC and ABC but noted that 'no animal species communicates in just this way'. However, more recent studies have shown limited combinatorial characteristics in vocal signalling of birds (Beer 1976) and primates (Robinson 1979, 1984; Cleveland & Snowdon 1982), and the black-capped chickadee, *Parus atricapillus*, has a manifestly combinatorial system like that envisioned by Wilson (Hailman et al. 1985). Furthermore, the chick-a-dee system is open; that is, it has no limit on the number of different call-types generated combinatorially from its repertoire of just four note-types, as is explained in more detail below. Still, no animal system has heretofore been shown to fit any objective definition of language offered by linguists. Here we show that the 'chick-a-dee' call-system of the black-capped chickadee has a computable syntax and thereby qualifies as a recursive language under an operational definition of structural linguistics.

Expressed in notation slightly modified from Gross (1972), languages have the following properties. There exists some finite vocabulary (V) of recurring signal-elements, such that V^* is the infinite set of all possible strings of concatenated elements of V. A language (R) consists of those strings actually used by the communicants, such that R is a proper subset of V^* (i.e. $R \subset V^*$ and $R \neq V^*$); that is, there exist strings in V^* that are not in the language R. The point of a rigorous structural definition of language is to show that there exists an explicit set of rules (syntax) by which the elements of V are combined to make communicable strings.

Languages form the following three-level hierarchy defined according to the degree to which their syntaxes are computable (Gross 1972). The arbitrary (or non-denumerable) languages cannot be described at all by Turing machines (Hopcroft 1984), which are hypothetical computers of the most basic sort (explained below). The 'recursively enumerable' languages are such that Turing machines can identify strings that are in the language but may continue computing indefinitely when encountering a string not in the language. The truly 'recursive' languages can be described entirely by a Turing machine, which is to say that every possible string can be identified explicitly as being either in the language or not.

A syntax is explicit (computable) if a Turing machine can be instructed to identify strings as being in the language or not (Gross 1972). The instructions (algorithm) are known as the 'characteristic function' of the syntax. The Turing machine reads input characters (I) one at a time, and according to the current state (S_j) of the machine it takes one specific action (α) and then changes to a new state (S_k) as instructed. (The new state may be the same as the original state S_j.) Instructions are therefore of the form (I, $S_j \rightarrow \alpha$, S_k). Permissible actions are of two types: either move the input tape one character (left or right), or write a new character over the old one. Some conceptions of Turing machines allow writing and moving the tape in one step (Hopcroft 1984), rendering a shorter list of instructions for a given task, but we follow here the machine conceived by Gross (1972). For purposes of linguistic evaluation, the Turing machine may write only # (blank: erasure), 1 (one: string in the language), or 0 (zero: string not in the language). Thus the instruction (A, $4 \rightarrow$L, 3) reads the input character A while in state S_4, moves the tape left one character and changes to state S_3. The set of instructions is defined (Gross 1972) as a characteristic function only if the Turing machine erases all input characters, writes a single 1 or 0 in place of the entire input string, and stops.

The black-capped chickadee utters about a dozen classes of vocalization, of which one class is the chick-a-dee call-system (Ficken et al. 1978a). Chick-a-dee calls are composed of four qualitatively different note-types uttered in strings of from one to at least 24 notes in length (Hailman et al. 1985). A spectrographed sample of 3479 calls contained 362 different call-types (i.e. different strings of note-types). Furthermore, it was shown mathematically that the call-system is open: that the repertoire of call-types is not restricted to the 362 found or any other finite number. The analysis even predicts the probabilities of occurrence of new call-types; for example, the probability of the 500th call-type is approximately 0·0001, meaning that a sample of 10 000 recorded calls will yield about 500 different call-types. The calls themselves are strung together in bouts (Ficken et al. 1978b), but these bouts consist of highly similar call-types and do not have the combinatorial character of calls themselves.

Note-types of chick-a-dee calls were arbitrarily designated A, B, C and D, and in 99·7% of calls the note-types occur in that order, where any note-type may be omitted, given once, or repeated a variable number of times. The very rare (0·3%) exceptional calls in the sample contain some permutation of order, such as the sequence -*BCB*-, which occurred within nine of the 11 exceptional calls. It is obvious intuitively that this syntax is computable, but we follow the formalism of Gross (1972) to demonstrate the point explicitly.

The input tape of chick-a-dee calls for a Turing machine therefore consists of the character set $\{A, B, C, D, \#\}$ where # is the silent period separating successive calls. The Turing machine begins in state S_0 and reads the first character in the input string. Moving the tape left (L) causes the next input character to the right in a string to be read. Table I is a set of 40 instructions making up a characteristic function of chick-a-dee calls. Rows of the matrix show the machine's state when reading an input character and columns show the character read; the cells dictate the action to be taken and the new state to be assumed. Thus when reading an A while in state S_0, the machine moves the input tape left (L) one character and changes to state S_1. The algorithm is such that the machine reads an entire string, then moves backward through the string erasing all input characters as it goes, until encountering the blank preceding the string. At this point if the machine is in state S_6 it writes a 0 or if in S_7 writes a 1, and then changes to state S_0. In either case it cannot read 0 or 1 in state S_0 so it stops (blocks), indicating that assessment is complete.

The characteristic function shown in the table is not necessarily the most parsimonious algorithm, but to show that chick-a-dee calls qualify as a recursive language parsimony is not required. The machine evaluates every possible input determinately (i.e. erases every possible finite string of input characters), writes either 1 or 0, and then stops. It writes 1 in place of every string in the form $A^aB^bC^cD^d$, where a, b, c and d are the lengths of substrings of component note-types and each has a zero or presumably finite value (e.g. $\infty > a \geqslant 0$), and writes 0 in place of any other string. The algorithm thus evaluates the extremely rare permuted strings as 'ungrammatical' (not in the language); it is possible that such strings were uttered by young birds that had not yet mastered the chick-a-dee syntax. The characteristic function of Table I also ignores the subtle tendency for each successive note in a repetitive substring to be of slightly lower acoustical frequency than the previous note, especially in repetition of A-notes. The communicative significance of this property (if any) is presently unknown.

This appears to be the first animal communication system shown to have a computable syntax. Furthermore, it is the only manifestly combinatorial system known so far in animals and the only

Table I. Matrix of instructions (characteristic function) for a determinate Turing machine to evaluate chick-a-dee calls*

	Input character (I)				
State (S_j)	A	B	C	D	#
0	L,1	L,2	L,3	L,4	L,0
1	L,1	L,2	L,3	L,4	R,7
2	L,5	L,2	L,3	L,4	R,7
3	L,5	L,5	L,3	L,4	R,7
4	L,5	L,5	L,5	L,4	R,7
5	L,5	L,5	L,5	L,5	R,6
6	#,5	#,5	#,5	#,5	0,0
7	#,1	#,2	#,3	#,4	1,0

* While in a specified state (rows) the machine reads a character (columns) from the input tape, and acts according to the instruction specified (cells). The act may be to move the tape left one character (L), right one character (R), erase the character read (#), write over the character read indicating that the string is not in the language (0), or write over indicating that the string is in the language (1). The machine then assumes the new state (0 through 7) indicated in the cell and reads the new input character now under its reading head. Permissible input characters are the four note-types of chick-a-dee calls (A, B, C and D) plus # (either the silence separating calls or an erasure by the machine). The machine begins in state S_0 and stops when it cannot act upon an input character (0 or 1 written by the machine).

system that is open (i.e. has no limit on the number of possible call-types that compute as grammatical). Some finite sets of behavioural sequences, such as nest-building in birds, may have computable rules, but such sequences are neither combinatorial nor open. Some song systems of oscine birds and some primate vocal systems have recurrence of elements, but these are not manifestly combinatorial since their diversity of signals depends upon large numbers of different sounds, not on combinations of a few different sounds; nor has it been shown that such systems are open. The joint occurrence of these three elements (combinatorial structure, openness and computable syntax) makes chick-a-dee calls far more like human language than any animal system yet described.

It seems unlikely that any system of animal communication will prove as complex as even the simplest human language. In the case of chick-a-dee calls we are still working to decode the information they contain. Our working hypothesis is that each of the four note-types refers to a qualitatively different referent about locomotory tendencies and repetitions refer to the relative strengths of those tendencies, so that every call-type has a unique meaning (Hailman et al. 1985; see also Hailman et al., in press). Not only would such a semantic system be far simpler than human language, but furthermore language has permutation of elements in addition to combination, and language has various levels of combination (phonemes combine to make morphemes, morphemes to make sentences, and so on). Nevertheless, the operational criterion of a computable syntax (along with combinatorial structure and openness) identifies animal systems likely to be of special interest in the quest to understand the evolution of human verbal communication.

The empirical work on chick-a-dee calling was done in collaboration with Robert W. Ficken, who helpfully criticized the manuscript. We are also grateful to Elizabeth D. Hailman and James Reinartz for aid and criticisms of various kinds, and to John G. Robinson, an anonymous reviewer and editor Charles T. Snowdon for comments on the manuscript. This is publication no. 78 of the University of Wisconsin-Milwaukee Field Station.

*JACK P. HAILMAN
†MILLICENT S. FICKEN

*Department of Zoology,
University of Wisconsin-Madison,
Madison, Wisconsin 53706, U.S.A.

†Department of Biological Sciences,
University of Wisconsin-Milwaukee,
Milwaukee, Wisconsin 53201, U.S.A.

References

Beer, C. G. 1976. Some complexities in the communication of gulls. *Ann. N.Y. Acad. Sci.*, **280**, 413–432.

Cleveland, J. & Snowdon, C. T. 1982. The complex vocal repertoire of the adult cotton-top tamarin *(Saguinus oedipus oedipus). Z. Tierpsychol.*, **58**, 231–270.

Ficken, M.S., Ficken, R.W. & Witkin, S.R. 1978a. Vocal repertoire of the black-capped chick-a-dee. *Auk*, **95**, 34–48.

Ficken, M.S., Hailman, J.P. & Ficken, R.W. 1978b. A model of repetitive behaviour illustrated by chick-a-dee calling. *Anim. Behav.*, **26**, 630–631.

Gross, M. 1972. *Mathematical Models in Linguistics.* Englewood Cliffs, New Jersey: Prentice Hall.

Hailman, J. P., Ficken, M. S. & Ficken, R. W. 1985. The 'chick-a-dee' calls of *Parus atricapillus*: a recombinant system of animal communication compared with written English. *Semiotica*, **56**, 191–224.

Hailman, J. P., Ficken, M. S. & Ficken, R. W. In press. Constraints on the structure of combinatorial 'chick-a-dee' calls. *Ethology*.

Hopcroft, J.E. 1984. Turing machines. *Scient. Am.*, **250**, 86–98.

Robinson, J. G. 1979. An analysis of the organization of vocal communication in the titi monkey *Callicebus moloch. Z. Tierpsychol.*, **49**, 381–405.

Robinson, J. G. 1984. Syntactic structures in the vocalizations of wedge-capped capuchin monkeys, *Cebus olivaceus. Behaviour*, **90**, 46–79.

Wilson, E. O. 1975. *Sociobiology: The New Synthesis.* Cambridge, Massachusetts: Harvard University Press.

(Received 10 March 1986; revised 14 June 1986; MS. number: AS-389)