

# Possible Neural Correlates of Talker Identification in Budgerigars

K. Mori

Research Institute of Logopedics and Phoniatrics, University of Tokyo  
7-3-1 Hongo, Bunkyo-ku, Tokyo 113 Japan

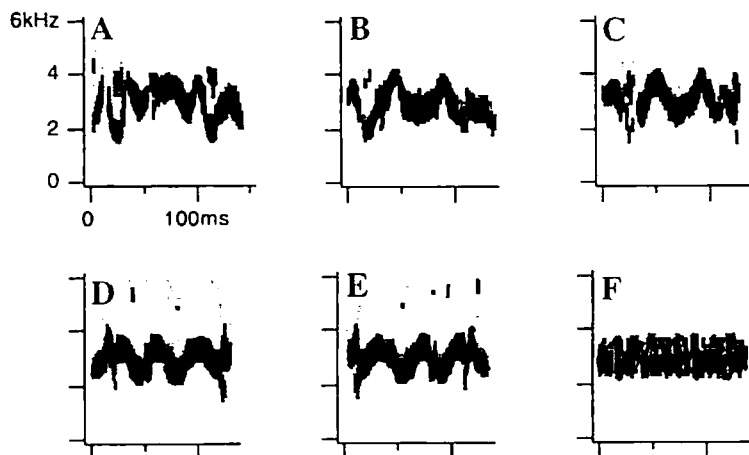
## Introduction

Birds are the only animals that can mimic human voice so well that humans take it for human voice. There are two orders of birds that are good mimics; passerine (West & King, 1990) and psittacine (Todt, 1975). Unlike other birds (Konishi, 1963; Kroodsma & Konishi, 1991), they fail to develop normal song or calls when they are raised without auditory feedback (Dooling, Gephart, Price, MacHale, & Brauth, 1987a; Konishi, 1965) or in isolation without a "tutor" (Dooling, Park, Brown, Okanoya, & Soli, 1987b). They memorize a tutor's voice and match their own vocal outputs with the memory. This matching process requires a mechanism for accurate sound recognition in the brain from which the vocal center derives information on how to modify vocalization. Vocal learning in passerines is usually limited to song, whereas parrots can learn widely different sounds. This psittacine ability is particularly interesting from the viewpoint of how sounds are represented in the brain. Also, songbirds learn song primarily in youth or critical periods, whereas parrots appear to be capable of vocal learning in adulthood as well, suggesting that both the auditory and vocal systems remain plastic.

Budgerigars (*Melopsittacus undulatus*) are obligatory social birds and vocal communication plays an important role in their social interactions. They have several distinct calls among which is a contact call, or distance call (Dooling, et al., 1987b). It is often used when they are out of sight from each other, thus helping to locate others. It is typically 120 to 200 ms in duration, 2 to 4 kHz in frequency and is characterized by rapid frequency modulations (fig. 1). Each bird has a stereotypic contact call. Because budgerigars learn vocalization from each other, the members of the same group tend to have similar contact calls (Brown, Dooling, & O'Grady, 1988). There is evidence that they perceive cagemate calls differently from those of non-cagemates (Brown, et al., 1988). Their average response latencies to the calls from different birds separated by cluster analysis only when they were tested to discriminate cagemate calls. Because they respond to their cagemate calls according to the callers, this may serve as a good animal model to study neural bases for talker identification.

In an attempt to investigate the neural correlates of this behavior, we studied the characteristics of neural responses to species specific calls and synthetic stimuli in the primary forebrain auditory area, field L and its adjacent structures. Neurons are found in the passerine brain that respond preferentially to the bird's own song (Doupe & Konishi, 1991; Margoliash, 1983) and species specific calls (Leppelsack & Vogt, 1976). This study is the first to show neuronal activity related to species specific calls in the psittacine forebrain. Part of the present work was presented elsewhere (Mori, 1992; Mori & Striedter, 1992; Mori & Striedter, 1993a; Mori & Striedter, 1993b).

Fig. 1. Sonograms of contact calls and modified calls; A, B and C, sonogram display of contact calls recorded from one bird over several months; D, sonogram of a contact call of a cagemate; E, reversal of D in the time domain; F, sonogram of band noise with the same overall spectral composition as D or E. Scales of A apply to all sonograms. A 3 ms time window was shifted by 2 ms steps for each time slice of the spectrograms.

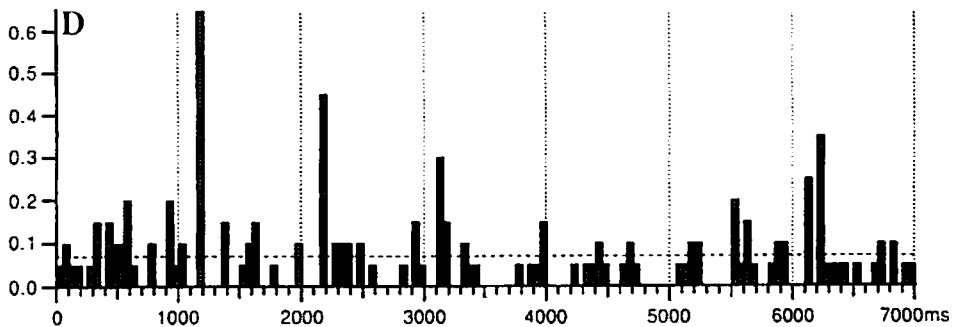
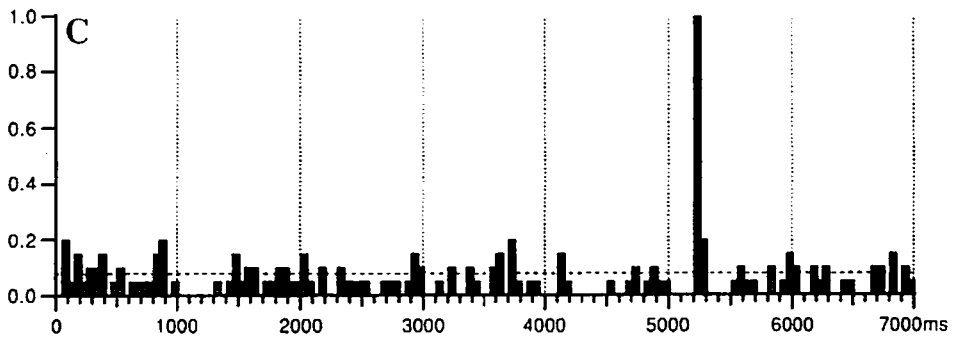
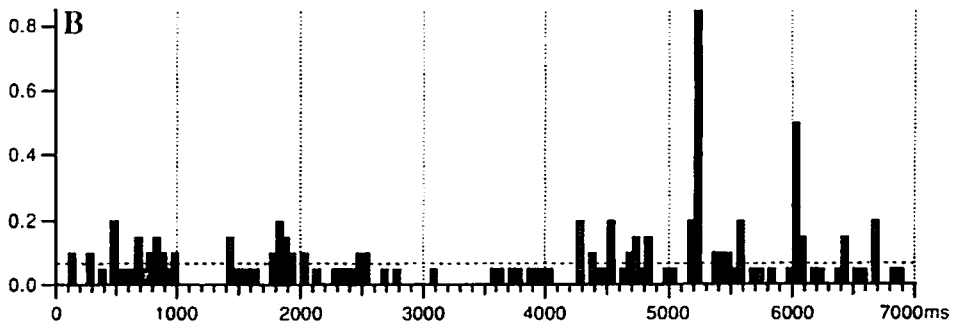
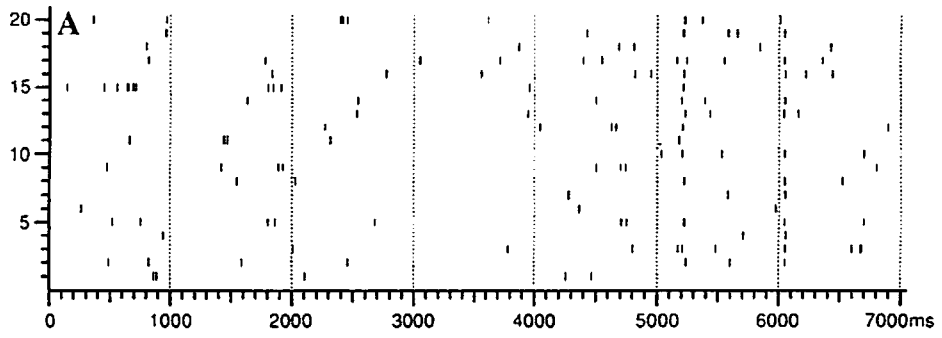


## Methods

Adult male budgerigars were obtained from a local dealer and housed together for several weeks before their calls were recorded in a sound attenuated box onto a tape. Calls were digitized off-line by a Macintosh IIfx computer with an A/D (analog to digital) converter (MacADIOS/II, GW Instruments) at the sampling rate of 20 kHz and stored on the hard disk (Mori, 1993). A wide band noise burst and tone bursts of 140 ms in duration with 10 ms of rise and fall, and upward and downward linear frequency sweeps ranging from 5 kHz/s to 320 kHz/s were also prepared and stored on the hard disk, as well as digitally modified calls. The modification methods of calls included (1) deleting the initial 25, 50 and 75 ms of the natural calls, (2) leaving the initial 25, 50, and 75 ms of the original calls and deleting the rest, (3) reversal of the calls in the time domain (fig. 1E), and (4) making band noise whose duration and overall frequency spectrum were same as the original calls (fig. 1F).

A stereotaxic head pin was affixed to the skull under Xylazine/Ketamine anesthesia several days prior to the recording session. Extracellular activity of single units was recorded stereotaxically in and around left field L with a parylene insulated tungsten microelectrode (125  $\mu\text{m}$  in diameter, A-M Systems, Inc.) under Urethane anesthesia (1.4 g/kg). Sound stimuli were converted from digital to analog (D/A), amplified by a McIntosh power amplifier, and delivered through a calibrated speaker in a sound attenuated box at 70 to 80 dB SPL at a rate of one per

Fig. 2 (next page). Responses of a single selective unit. A raster display of spikes as vertical bars (A) and a peri-stimulus time histogram (PSTH) with 50 ms bins (B) show the responses of a unit to white noise played back at 1 s from the beginning of each cycle, contact calls of four different birds at 2 through 5 s, and one thwart call at 6 s. The best response was elicited by the contact call at 5 s (the best call). There were 20 cycles of recording with the same stimulus set as shown in A. C shows a PSTH of responses to band noise at 1 s, call reversal at 3 s and the best call played forward at 5 s, are shown as a PSTH. D shows a PSTH of responses to abridged versions of the best call. The first three calls at 1, 2 and 3 s were cut by 25, 50 and 75 ms from the beginning (see Methods, modification 1). The last three at 4, 5, and 6 s were only the initial 25, 50 and 75 ms (modification 2). Calls were played at the beginning of every second. The ordinates for B through D are average spike count per stimulus cycle. The horizontal dotted line shows an averaged spontaneous activity. The first second was for spontaneous activity without sound. There was at least 3 s of inter-cycle interval.



second, unless otherwise noted. Each set of stimuli was presented ten or more times, a response spike histogram was constructed and responses to respective stimuli were statistically tested. Stimuli consisting of wide band noise and various calls were used as search stimuli. Once units were isolated, their preference for noise or calls was determined. Units preferring calls to noise were further tested with additional contact calls, tone bursts, frequency sweeps and various modified calls. Those units responding to only one contact call significantly better than to the remaining calls or white noise were classified as "call selective units", and were tested with various modified versions of the best call. Electrolytic lesions were made and recording tracks were later reconstructed histologically.

## Results

Figure 1 shows some of the contact calls and modified calls used in the present study as frequency-time sound spectrograms (sonagrams). Figure 1A through C shows changes in the contact call of one bird after repeated association with a new group of birds over a period of several months. Figure 1D shows the sonagram of a cagemate call when the call shown in fig. 1C was recorded. It is more similar to C than A or B, suggesting a mutual learning between those cagemates.

Among 200 isolated single units that responded to one or more of the search stimuli, 55 % of them responded significantly better to one or more calls than to noise. Thirty percent of them did not show significant preference. Only 15% preferred noise, of which 40 % did not respond to any tone burst.

Of the call preferring units, 86% responded to tone bursts of various frequencies, while 44% of them did not respond to white noise, which suggests that there is inhibitory interaction among different frequencies. Ninety-four percent of the call preferring units responded to at least one upward or downward frequency sweep.

Nine percent of the call preferring units responded predominantly to only one of the contact calls in our stimulus set (call selective units), although sonagrams of some of the calls were similar. Responses of a typical selective unit are shown in Fig. 2. This unit responded with a long latency to the call played at 5 s from the start of each cycle (the best call, which was the one shown in fig. 1C). It did not respond to white noise played at 1 s, or rather it did so with possible suppression of spikes. Sixty-seven percent of the selective units still responded to frequency sweeps and tone bursts, but there were units that responded to none of these. When selective units were tested with their best call played forward and backward, five out of nine tested units showed a significantly better response to the forward call. They also responded better to the forward call than to the band noise whose duration and overall spectrum were same as the call (fig. 2C). The other selective units were less discriminate between forward and reverse calls. When the selective units were tested with abridged calls, some showed graded responses, as observed in behavioral discrimination tests of similarly abridged contact calls (Park & Dooling, 1986). An

example of this graded response is shown in fig. 2D. The longer the initial deletion was, the smaller the response became, as seen at 1 through 3 s in the figure. The initial 75 ms call segment elicited a small response, but the segment of only 50 or 25 ms did not.

Among the call preferring units, 4 % were located in hyperstriatum ventrale (HV), which is dorsal to field L, 67 % in L1, 13 % in L2, 15 % in L3, and 1 % in paleoarchistriatum. The response characteristics of HV units were indistinguishable from those of the field L neurons. Paleoarchistriatum units usually showed regular spontaneous activity and many of them responded to sound stimuli by decreasing the firing rate.

By contrast, call selective units were missing from L2, although two units were located on the border of L2 with either L1 or L3. The remaining selective units were found mostly in L1 and L3, and a few in HV.

## Discussion

Of the noise preferring units, 40 % did not respond to tone bursts, which suggests that their response is not a simple sum of various frequency inputs, but that there is a more complex tuning mechanism involved. These complex neurons may participate in detecting calls such as alarm calls which contain a broader range of frequencies than do contact calls.

Eighty-six percent of the call preferring units responded to tone bursts. Although it is possible that simple frequency tuned units could respond better to calls than to noise because of the greater spectral concentration in calls, forty-five percent of those call preferring units did not respond to noise at all. This suggests that call preference involves processing more complex than simple frequency tuning, e.g., inter-frequency inhibition. This also validates our strategy to include not only white noise but also various calls as search stimuli.

More than half of the units in the present study responded better to a call than to white noise, but only 15% preferred noise. This is in contrast to a study of Guinea fowl field L (Scheich, Langner, & Bonke, 1979), where only 53 % of neurons responded to the species specific calls, although majority of those neurons responded to tone bursts as well. This discrepancy can at least partly be explained by the fact that guinea fowl's vocalization is spectrally more complex than budgerigar's contact calls, provided that many neurons in the Guinea fowl field L also have interfrequency inhibition (Scheich & Bonke, 1981). A study with starling field L (Müller & Leppelsack, 1985) also shows that many units responded to tone burst stimuli, which suggests that the responses of field L neurons to tone bursts are similar among different species.

As much as 94% of the call preferring units in budgerigars responded to frequency sweeps. Although budgerigar calls are abundant in frequency sweeps, this may not necessarily imply adaptation to species specific calls in field L, because FM sweeps excite various units including simple frequency tuned neurons. However, data by the [<sup>14</sup>C]2-deoxyglucose method showed that FM sweeps produced excitation in larger areas than tone bursts in field L of some species including the budgerigar, but not others (Müller & Scheich, 1985; Scheich & Bonke, 1981). It

remains to be seen if the responsiveness to FM sweeps of field L units is correlated, across species, to the amount of FM components in species specific sounds.

Field L is divided into 3 layers, dorsal L1, middle L2 and ventral L3 (Müller & Scheich, 1985). L2 is the main input layer of field L from thalamic nucleus ovoidalis (Bonke, Scheich, & Langner, 1979). In our study, even though there were more L1 units than L2 or L3, which may reflect our sampling bias towards anterior field L than posterior, we sampled a fair number of L2 units. However, we did not observe any selective units in L2, except that there were two selective units on the border with L1 and L3 respectively. This tendency is similar to that observed in field L of the Guinea fowl (Bonke, et al., 1979) and the starling (Müller & Leppelsack, 1985) and may imply a general rule of hierarchical signal processing in field L, with L1 and L3 processing higher order information than L2.

We found call selective units in field L and HV. The majority of these decreased activity when their best call was reversed in the time domain. This corresponds well to the behavioral data that cagemate birds failed to classify calls according to callers when calls were reversed (Brown, et al., 1988). Some of the selective units showed graded responses to abridged calls, very similarly to behavioral performance with degraded calls (Park & Dooling, 1986). If field L is responsible for talker identification behavior, it must be possible to explain the birds' behavior in terms of the neural response properties of field L. Even if it is not, there may be neurons in field L that feed necessary information to other nuclei that are responsible for talker identification. Since our data are consistent with behavioral data, it is possible that field L is involved in the neural processing of talker identification behavior in the budgerigar. However, other means than unit recording as employed in this study are also necessary to determine how and in what level of processing field L is involved in the behavior.

Since we found units highly selective to single calls, it is unlikely that the neural representation of the outer world in field L is exclusively distributed. Our data are more compatible with a sparse population coding system or with a "grandmother" hypothesis. Because there is a whole range of selectivity among neurons in field L, it is possible that, when birds learn new cagemate calls, selective neurons to the new calls emerge from less selective units. Similar plasticity is already found elsewhere (Miyashita, Chang, & Mori, 1992). However, our present data do not specifically address these points. Because budgerigars seem to be able to learn to discriminate individuals by hearing calls, they would make a good animal model to investigate the neural representation of talkers in the brain as well as the mechanism for auditory learning.

## Acknowledgments

This work was supported by Ministry of Education, Science and Culture of Japan and Long-term fellowship from Human Frontier Science Organization. The author is grateful to Masakazu Konishi and Georg Strieder for discussion and comments.

## References

Bonke, D., Scheich, H. & 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. *Journal of Comparative Physiology A*, **132**, 243-255.

Brown, S. D., Dooling, R. J. & O'Grady, K. (1988). Perceptual organization of acoustic stimuli by budgerigars (*Melopsittacus undulatus*): III. Contact calls. *Journal of Comparative Psychology*, **102**(3), 236-247.

Dooling, R. J., Gephart, B. F., Price, P. H., MacHale, C. & Brauth, S. E. (1987a). Effects of Deafening on the contact call of the budgerigar, *Melopsittacus undulatus*. *Animal Behaviour*, **35**(4), 1264-1266.

Dooling, R. J., Park, T. J., Brown, S. D., Okanoya, K. & Soli, S. D. (1987b). Perceptual organization of acoustic stimuli by budgerigars (*Melopsittacus undulatus*): II. Vocal signals. *Journal of Comparative Psychology*, **101**(4), 367-381.

Doupe, A. J. & Konishi, M. (1991). Song-selective auditory circuits in the vocal control system of the zebra finch. *Proceedings of National Academy of Science USA*, **88**, 11339-11343.

Konishi, M. (1963). The role of auditory feedback in the vocal behavior of the domestic fowl. *Zeitschrift fur Tierpsychologie*, **20**, 349-367.

Konishi, M. (1965). The role of auditory feedback in the control of vocalization in the white-crowned sparrow. *Zeitschrift fur Tierpsychologie*, **22**, 770-783.

Kroodsma, D. E. & Konishi, M. (1991). A suboscine bird (eastern phoebe, *Sayornis phoebe*) develops normal song without auditory feedback. *Animal Behaviour*, **42**, 477-487.

Leppelsack, H. J. & Vogt, M. (1976). Responses of auditory neurons in the forebrain of a songbird to stimulation with species-specific sounds. *Journal of Comparative Physiology*, **107**, 263-274.

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

Miyashita, Y., Chang, H. S. & Mori, K. (1992). A hypothesis on primal long-term memory: Neurophysiological evidence in the primate temporal cortex. In L. R. Squire & N. Butters (Eds.), *Neuropsychology of Memory* (pp. 347-356). New York: The Guilford Press.

Mori, K. (1992). Response characteristics of field L neurons to species specific calls of budgerigars. In *The Research Symposium "Neuroethological studies of vocal communication"*. (p. 1 in Japanese). National Institute for Physiological Sciences, Okazaki, Japan.

Mori, K. (1993). A personal computer based environment for neurophysiological experiments that provides easy on-line modification tools of auditory stimuli. *Otology Japan*, **3**, 75 in Japanese.

Mori, K. & Striedter, G. (1992). Neurons in field L of Budgerigars prefer species specific calls to white noise. *Society for Neuroscience Abstracts*, **18**, 527.

Mori, K. & Striedter, G. (1993a). Possible neural basis of talker identification in budgerigars. In *International Symposium on Speech Production and Communication Disorders*. University of Tokyo, Tokyo, Japan.

Mori, K. & Striedter, G. (1993b). Possible neural correlates of call discrimination in the budgerigar forebrain. In *Seminar Neurobiology V - From molecule to Brain*, (pp. 17-18). Tokyo, Japan. Japan Intractable Disease Research Foundation.

Müller, C. M. & Leppelsack, H.-J. (1985). Feature extraction and tonotopic organization in the avian auditory forebrain. *Experimental Brain Research*, **59**, 587-599.

Müller, S. C. & Scheich, H. (1985). Functional organization of the avian auditory field L. A comparative 2DG study. *Journal of Comparative Physiology A*, **156**, 1-12.

Park, T. J. & Dooling, R. J. (1986). Perception of degraded vocalizations by budgerigars (*Melopsittacus undulatus*). *Animal Learning & Behavior*, **14**(4), 359-364.

Scheich, H. & Bonke, B. A. (1981). Tone- versus FM-incuded patterns of excitation and suppression in the 14-C-2-deoxyglucose labeled auditory "cortex" of the Guinea fowl. *Experimental Brain Research*, **44**, 445-449.

Scheich, H., Langner, G. & 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 lambus-like calls. *Journal of Comparative Physiology A*, **132**, 257-276.

Todt, D. (1975). Social learning of vocal patterns and modes of their application in Grey parrots (*Psittacus erithacus*). *Zeitschrift fur Tierpsychologie*, **39**, 178-188.

West, M. J. & King, A. P. (1990). Mozart's Starling. *American Scientist*, **78**, 106-114.