100. Lilly, John C. 1967. Dolphin's Mimicry as a Unique Ability and a Step Towards Understanding in Research in Verbal Behavior and Some Neurophysiological Implications. Kurt Salzinger and Suzanne Salzinger, Eds. Conference on Verbal Behavior, N.Y.C. 1965. Academic Press, New York City. P. 21-27

8

DS

REPRINTED FROM
RESEARCH IN VERBAL BEHAVIOR
AND SOME NEUROPHYSIOLOGICAL IMPLICATIONS

ACADEMIC PRESS INC., NEW YORK

1 ord
95

Dolphin's Vocal Mimicry as a Unique Ability and a Step Toward Understanding [1]

John C. Lilly

Communication Research Institute, Miami, Florida

In our program of human-dolphin communication, we are finding that the human limitations are as controlling as the dolphin limitations. I am making a plea for those who work with the humans to give us a far better model of Man the Communicator than we have. Present models of the human's language capabilities and its acquisition are quite deficient. We fall out here at a very deep level. Until we have such a model it is difficult to model the dolphin's communication.

I came upon the dolphin's vocal mimicry of the human voice by accident in 1957, while pursuing some neurophysiological aims with the dolphin. It took us some time to believe that it was happening because we had not read the previous literature, namely, that of Aristotle (384–322 B.C.). He is the author who first reported "a voice like that of the human" from the dolphin in air. It is nice to have a gap of only twenty-three centuries in the literature, but it is rather disheartening to find a "new" phenomenon of this sort, and then learn that the previous observer's observations were discredited for 2300 years.

Nobody could have convinced me before we did these experiments that a small whale (*Tursiops truncatus*) would attempt to mimic human speech (or any other noises) in air or that mimicry was one of its abilities. As I recounted in the book *Man and Dolphin* we discovered that he emitted these sounds when we failed to give an expected reward. (We realized later that this was a use of vocalization as an operant demand on a responding environment.) The first noise that the dolphin reproduced was an electronic whistle that was going on in the background (we were tuning some apparatus), and he began to match frequencies. They were rather high frequencies.

Somebody had laughed in the laboratory and he made some rhythmic

[1] This investigation was supported in part by Public Health Service Research Grant NB–03097 from the National Institute of Neurological Disease and Blindness by Grant K6–MH–18,700 from the National Institute of Mental Health; and by the United States Air Force, Air Force Office of Scientific Research Grant No. AF-AFOSR 65–65.

harsh noises that, in a gross sort of way, resembled this laughter. We noticed that, when I was dictating notes to the tape recorder, some physical aspects of what I was saying were being reproduced by the animal.

We did not pick this up immediately. We had to slow our tapes down to hear the mimicry and then speed them up again as one does to understand synthetic speech from a computer. (It helps to do this so that one may take the elements apart and then put them back together again at normal speed.) Listening to these recordings, certain other people did not hear mimicry; these first tapes were not convincing to people who lacked an ear for timing, rhythm, and accent. We found that people who knew foreign languages and speech therapists, both of whom were good phoneticians, could pick up the mimicry immediately; but that zoologists, marine biologists, ethologists, and neurophysiologists who had not paid very close attention to the way things really sound did not pick it up. In other words, with trained ears one could hear the phenomenon, and if one were trained with the dolphins themselves one could do so even more easily. This necessity for trained human listeners has been one of our major stumbling blocks.

We do not have an advantage that one has with the mynah bird: that when he copies, the mynah makes a good copy. The dolphin does not copy the sounds well, but he does a better job than the mynah with timing, number of bursts, and rhythm; he locks in with a human in high-speed exchanges (Lilly, 1965).

To continue the history, we found in 1958 and 1959 that we could elicit the mimicry by simpler methods. We went into the dolphin's tank, started shouting at him, and he "shouted" back in air. My assistant (A. M. Miller) and I had been discussing our approaches one day, and she said, "All right let's go in and try talking with Elvar." We went in and she talked loudly at Elvar. He then raised his blowhole out of the water and emitted mimicking noises. It was not until we wanted him to do this and set the conditions (almost unconsciously) correctly to get it, that we finally got it. In other words, we had freed ourselves up from the original conditions of the discovery. We found that one organism stimulating the other organism with sound (the human and dolphin) worked.

We went on with the usual operant conditioning until we began to realize that we did not know what were the reinforcements for the dolphin. We demonstrated to our satisfaction that the dolphin operates very well at first with food reward, but that soon he does not need it. We started new dolphins with food reward, but at some point along the way, the dolphin abandons the need for food reward.

Two years ago we began to use nonsense syllables. We did this to avoid the trap of attaching meaning to the dolphin's replies. Our plan was to investigate the vocal capability of the dolphin in some *physically specifiable aspects* of vocal mimicry. We started out with 196 separate vowel-consonant, consonant-vowel combinations; made up randomized lists; and tested a number of responses. We first trained a dolphin on the list itself; we used the vowel-consonant and consonant-vowel combinations and forced him to practice his transaction control. He mastered this difficult task: he gave back only two pairs of sounds (no matter what they sounded like) for two pairs of sounds from us.

We tested his ability to do this under other conditions. The vowel-consonants and consonant-vowels were divided into groups, and presentations from one to ten in each group were given him.

To illustrate, we would say "no, ot, toe," a group of three. If we then obtained from the dolphin a group of three sounds, we accepted that as correct performance from the dolphin, irrespective of what it sounded like. (We did not go into the "sounding like" at this point). We just recorded what happened. The first dolphin caught onto this in one presentation of 194 items given in a period of 20 minutes; his errors were 36%. On the second presentation 2 hours later he had 92% correct (8% errors). In this case he responded steadily over a 12-minute period with sonic bursts coming at the rate of one every 0.6 second.

The dolphin's delay in answering after each presentation varied from 0.1 second to 0.5 second. The peak of the distribution curve is about 0.35 second, which is just about the human delay between the items in the list. That is, in a humanly emitted group such as "one, two, three, four, five," the silences between the sonic bursts average something on the order of 0.2 to 0.4 second depending on how the list is given. He matched burst durations, interburst silence durations; his latency (human end of list to dolphin beginning) was comparable to the interburst silence duration, and sometimes it was even less. Therefore, we concluded that he was using vocal clues from the terminal item in the list. The fact that he was able to key himself off as one finished the list meant that we were somehow furnishing vocal clues that said, in effect, "This is the last item in the list"; he was picking up these clues.

As controls, we put the items onto a tape to rid him of the presence of the human operator (cf. "clever Hans," the counting horse case). We put in a machine feeder for the food reward. (We were still working on a one-to-one reinforcement schedule at this point). We found no deterioration of performance. In other words, we could program him with a tape that was recorded by someone else at a different time, and he would perform just as with the operator in the room. We did all the necessary controls for the "clever Hans" effect. The operator was removed completely. We programmed the experiment from tape; the dolphin could see no humans, could hear no real operators, and could get no clues other than those on tape. We eliminated all clicks and other clues that the dolphin might be able to use.

Thus we achieved a performance with the dolphin that has not been achieved with the "talking" birds. Humans can do these tasks; dolphins can do them; birds cannot. We chose ten nonsense syllables as a maximum based on the results in the literature with humans. The dolphin's errors are just as great at three as they are at ten; that is, there is a randomly distributed group of errors up to and including ten items. His capability at ten was quite as good as it was at three or five.

To get back to the problem of reinforcement, the beginning list was used with immediate one-to-one reinforcement. Later, we fed the dolphin to satiation before the experiments. He accepted the fish in his mouth after each successful playback, and dropped it to the bottom of his tank. He accumulated a pile of fish under his beak and, at the end of the experiment, he handed the last fish back to the experimenter. Now, at times we use no food reinforcement. We dump all the fish into the tank and let the dolphin eat his fill. We do not know what all the other reinforcers are. It may be the activity itself; it may be the transaction with the human; or it may be other variables of which we are not yet aware.

Let us compare a tape recorder, a mynah bird, a dolphin, and a human in terms of the size of the computer available and in terms of input-outputs, storage and computations available. Vocal mimicry can be modeled in terms of programming. The organism hears A; he then makes a sound B and hears B, and centrally he compares A to B. This computation generates an error function with a servo-feedback. The CNS thus makes an analysis and initiates a new output. The aim is the reduction of the error—in other words, to increase the correspondence between A and B. There is nothing new in this. It is quite an old model. It has been used for gun pointing, for example. I think it is a useful model here, because this is just what the dolphins will do—correct themselves at each new presentation. The reply emissions from repeats of the same sounds by the human show progressive changes.

A dolphin is not bored with this kind of work if one does not go beyond 15-minute periods. A bright child or adult human does best with similar periods. A dolphin works solidly with you for about 15–20 minutes, then he'll swim away, take a rest of 5–20 minutes, and come back. He'll do this all day long, as long as you're willing to go along with it.

He reduces his errors. If the operator gives him a sequence in which the human says "ball," the dolphin comes back with a sound remotely like "ball," a high pitched sound. Then if the operator says "no, say, *ball,*" the dolphin does not copy "no, say"—he copies only the word "ball," and he gives a *new* copy better than the previous one. Thus, he will vary what he has said in respect to what the human said. But he selects out of all that the human says only that which the human says to say. He apparently learns the metalanguage used for giving him directions.

Sonographically the second human repeat of the word "ball" is not identical with the first one. There seems to be a novelty effect. The dolphin tries to approximate this new thing that he has heard. Even though one may think one is saying the same word, one is really not pronouncing it exactly the same way. Any of the experts on speech can tell you that it is very hard to give a long series of the same word and not have the later versions vary considerably. This novelty itself seems to be reinforcing to the dolphin.

If one gives the dolphin the same list from a tape again and again and again, so that he has the same set of stimuli, he will work for a few sessions and then drop it completely. You can't get him back to that same group. But if one then varies something in the setup such as changing to the whispered voice from the voiced emission he will take up the same list again. If one can add *acoustic novelty,* something new to work on, at each presentation he will then keep going.

How many dolphins will do this? We spent five years on the first two; we have spent about four years on the third, and the fourth has been working steadily for the last six months with Miss Howe in Saint Thomas. The fourth dolphin has shaped up better than any of the others to date.

In the fourth case we are doing some experiments which I suggested in *Man and Dolphin,* of the mother-child kind of situation. Miss Howe lived with Peter day and night 24 hours a day for 6 days a week for 6 months. She did a preliminary series of 1 week, 7 days and nights. She slept in the same room with the dolphin; the room was flooded to about sixteen inches with sea water. Dolphin sleeping habits can be rough on a human. Since dolphins take little cat naps all during the day and all during the night, they may want to work or play at 4:00 A.M. just as well as at 4:00 P.M. However, this technique has been giving us new information.

This dolphin is learning rapidly; he is beginning to combine vocalized words, objects, actions, and people. When he wants Margaret he makes the same kind of sound to fetch her from a far room as he did in many previous instances. We are getting a certain amount of reproducibility and stability in terms of his reproduction of words to get a response from the environment.

We have a "free floating" environment and a "free floating" response situation which seems to be very powerful for learning and teaching. Although we found that the formal reinforcements give machine-like precision in shaping a dolphin up, and he will give a magnificent performance, we feel that reinforcement theory as presently enunciated, and as understood by us, needs elaboration. I am proposing that we go to the computer analog:

programming, metaprogramming, subroutines, and all of this, and see whether this will not work better with these large-brained organisms.

Their huge computer-brain (it is larger than ours) is much more devoted to acoustic computations than ours. Our neurological group has been carefully examining this brain. Their visual input fibers are one-tenth the number of ours; we have a million, they have a hundred thousand. Their acoustic fibers are two and a half times the number of ours, on the order of one hundred and fifty thousand.

Their phonation apparatus is completely innervated by the facial nerve, which has 30,000 fibers in it. The sum of the fibers leading to our phonation apparatus has about the same number. Therefore, on the output side their control mechanism has about the same number of channels as ours.

On the input-acoustic side some of the fibers are given over to a sonar function; they can use the sonar function for communication also. They use it as a tight link when they want secrecy and do not want humans or sharks or something else to hear them. We have heard (on proper radio receiver and hydrophones) dolphins locked in at 150 kilocycles communicating rapidly back and forth with very formal transactions going on completely beyond our hearing.

I did not hear anything about the physical variables of the mynah bird, and I would be fascinated to learn what they are. I want to put in a bit of warning here. The ultrasonics are important in most species in which they have been investigated. As you now know, most of the rodents apparently are working at very high frequencies (like the bats), and there is apparently sonar even in the shrew. I was just wondering whether the mynah bird may not have a sonar also.

We must remember also that we put out ultrasonics. If you listen with proper equipment when a person is speaking, you can hear quite a lot of noise around 50 kilocycles. This could be very confusing to an animal that has a high sensitivity in that region. The dolphin's output runs from a minimum of about 400 cycles. His first formant in mimicking us is around 400 cycles. His upper formants reach 7 or 8 kc: too high for our ears. We are devising equipment to bring all this down into a more usable range for the human ear. Within the above limits the dolphin uses at least four formants. The upper end of communication dolphin to dolphin is of the order of 20 kc. His sonar begins at about 30 kc and goes on to about 250 kc, with a peak somewhere around 100–120. This makes the physical situation difficult for both of us.

In addition, below the blowhole the dolphin has a double phonation apparatus; each nasal passageway is under independent voluntary control. High speed motion pictures with X-rays and with visible light during the phonation activity show that one side can operate alone, then the other side alone, or a coordinated operation takes place involving the two sides.

We put hydrophones on the head of the dolphin out of water and did a two-channel recording of both sides separately. We then listened binaurally "stereo." We heard sounds on the left and then on the right and then sounds moving from left to right. We are calling this "double phonation" and "stereophonation" to separate these two processes.

If you examine the dolphin's phonation apparatus, you can see that he has the possibility of coupling the two sides through the nasal membrane and through the two diagonal membranes. These membranes are his main phonators and are the analog of our vocal cords. One sees in the literature the statement, "Whales can't talk, they don't have vocal cords." But they do have diagonal membranes and they have nasal "tongues," the so-called "nasal plug" which is analogous to our tongue. If one puts one's finger down in one side of the blowhole one can feel the dolphin using one of his two tongue-plugs in the way we use our tongue, trying to force the finger out of the blowhole. He also has to breathe through this mechanism; his speech is interrupted during the taking of a breath. He uses mainly the right side for respiration. He can continue to vocalize on the left side during a respiration if he is forced to. We have forced him to do this. When he is copying us he will make certain tones on the right and other tones on the left in a coordinated, locked-in fashion. This is a sophisticated kind of vocal performance.

It does not look as though he has cerebral dominance in the sense that we do. It looks as though he has "alternating dominance" and a "coordinated dominence" between the two sides. The only midline structure he has for his vocalization activity is the sonar apparatus, which we now pin down to being the larynx. It is very different and quite separate from the communication at the lower frequencies.

This is a summation of several years of research with the dolphin–human communication problems. We are pursuing this strange and new field with new instruments and new methods. We need bright and flexible help from many fields including human acoustics and speech, psychology, computers, the humanities, psychoanalysis, psychopharmacology, veterinary medicine. There is a big future here; I hope Man sees and seizes this opportunity for new vistas, new thinking, new philosophies, and a new breakthrough to escape his solipsistic preoccupation and anthropocentric and anthropomorphic self-adulation of himself and of his fellows: the dolphins are still for us and with us. We need them.

Group Discussion

Staats: I do not easily cast aside the learning principles in dealing with these topics. Learning theory becomes more powerful when one considers conditioned reinforcements as well as primary reinforcements. The behavior of Dr. Lilly's dolphins can be considered in terms of learning: for example, that making matching (imitational) sounds becomes reinforcing for dolphins. That is, with the dolphin, there is the possibility that in its natural environment much reinforcement is forthcoming for vocal imitational behavior in terms of his sonar. When the dolphin makes a vocal sound and gets a repetition of it back (as an echo) the consequences contingent upon this are, for example, that he avoids striking things, he finds food, etc., consequences having reinforcing value.

Lilly: Because the dolphin is a big-brained animal with better acoustic analyzers than ours, but in a strange frequency region, reinforcement theory has to be very subtly applied. Strict reinforcement contingencies for vocal behavior are not so obvious, and generalization does take place so that the organism obtains additional reinforcements from other sources, for example, skin stimulation.

Geschwind: If an organism depends on secondary reinforcement alone, even in an animal like the chimpanzee, the behavior very rapidly fails. If the dolphin does depend so heavily on secondary reinforcers, then there is something quite distinctive about him, just as in humans.