



On Systems Communicability

Shelia Guberman

1. Keldysh Institut of Applied Mathematics

Abstract: The article examines the fundamental mechanisms that make communication between living beings possible, focusing primarily on human interaction through acoustic and visual channels based on physical activity. It argues that any effective communication system must rely on signals that can be generated and perceived by all members of a species, which requires that communication alphabets be defined in relative, rather than absolute, physical terms. Using speech as a central example, the article discusses how phonemes are encoded through articulatory movements and proposes that perception occurs through internal reconstruction of these movements. This view aligns with the motor theory of perception, according to which understanding speech involves activating the same motor patterns used to produce it. Extending this principle, the paper suggests that recognition in other domains—such as handwriting, visual forms, and schematic drawings—can also be understood as the reconstruction or imitation of the processes that generated the observed objects. Handwritten letters are interpreted as traces of pen movements rather than static images, and geometric forms can be recognized by modeling their generative construction. More broadly, perception in music, dance, and sports is linked to internal imitation of movements, with richer motor experience leading to deeper and more emotional understanding. Farther the article explores communication in animals, emphasizing that differences in physical organization lead to different communication languages and perceptual worlds. Mutual understanding across species is therefore limited unless signals are adapted to the motor capabilities of each species. The article concludes that imitation and the formation of connections between motor acts and sensory patterns play a central role in learning, perception, and communication across domains.

INTRODUCTION

The article discusses the question of how the communication function between living beings is realized. Generally, we will deal with communication issues between people and only to a small degree will we be concerned with communication of representatives of other species among themselves and between them and people. For communication implementation between individuals of one species, it is necessary that each individual separately possess the ability to generate necessary signals (the alphabet of communication language). This requirement isn't as trivial as it can seem at first sight as it imposes essential restrictions on a choice of the alphabet of communication language. As any pair of individuals of one species differs in proportions, and in particular in size, elements of the alphabet have to be described not in absolute values, but in relative terms. For example, it is impossible to arrange a naval semaphore flag code which distinguishes letters by the height of raising the flag over the earth because the short signaler can't lift a flag as high as the tall one. Therefore, signs of a flag code are described in terms of angles between the hand and a vertical line: in this language the signaler of any stature can convey a message. Similarly, the speech code can't be based on the distinction of signals on the absolute frequencies of a sound, because very low frequency sounds aren't available to the child, and to a bass aren't

available the high frequency sounds. The alphabet of a speech code uses terms like *occlusion between tongue and teeth* or *lips closing*, etc. Such signals are available to any person.

We can now define the limits of this paper. We won't be interested in ways of communication such as chemical (smells) and temperature change, but we will be limited to acoustic and visual communication channels. Moreover, we will be limited to ways of communication by means of physical activity (for example, the above-mentioned flag and articulation codes) and therefore we won't be concerned, for example, with color codes. Let us consider some principles of speech communication. The alphabet of a speech code is made up of phonemes. Each phoneme is defined by a set of values of some parameters of the articulation apparatus: existence of an occlusion or bows, place of articulation, vibrations of vocal chords, position of tongue etc. The perception of a speech code becomes complicated because the transfer of the message is accompanied by a lack of direct visibility (and it is the most important advantage of voice channel of communication). Articulatory movements are transformed into a complex mix of vibrations of the air. The listener's task is to transfer oscillating air pressure perceived by the ear into an initial articulation code. It appears that it can be done. The existence or absence of basic tone (vibration of vocal cord) is possible by rough spectral selection: the frequency of vocal cord vibration is always less than 600 Hz. The existence of occlusion (in stop consonants) is determined by a sharp increase in the energy of a sound. The directions of moving of the first and second formant (up or down) depend on the place of articulation. . A small opening in the articulatory tract is recognized by a characteristic noise. When perceiving speech we perceive it directly in an articulation code.

MOTOR THEORY OF SPEECH PERCEPTION

Such a model of perception (which is called motor theory) appeared in psychology a couple of times. Mach was the first to suggest that in perception of rhythm there are two ingredients: tone-sensation and muscle-sensation. Ehrenfels in 1890 followed him: "Each perceived step from note to note caused in us a characteristic sensation (or feeling) belonging not to the sense of sound but to some other sphere (perhaps involving nervous or muscular sensations)" [Ehrenfels 1890]. In 1905 R. Stetson [1] explained the motor theory of rhythm perception: "It is not necessary that joints be involved, but according to changes in muscular conditions which stand in consciousness as movements are according to Stetson essential to any rhythm, whether "perceived" or "produced"" In 1957 Lieberman [2] revived the motor theory of speech perception. The theory claimed that people perceive spoken words by identifying the vocal tract gestures with which they are pronounced. The scientific community once more rejected the theory. 40 years later after discovery of mirror neurons on the wave of simulation theory D'Ausilio et al. [3] proposed "a modified motor theory of speech perception according to which speech comprehension is grounded in motor circuits". The conclusion is based on neurological experiments on the lip or tongue areas of M1 of the human motor cortex.

Directly perceiving speech in its original alphabet develops during speech acquisition. Schematically it can be presented as follows. The child produces various sounds, i.e. makes various movements with his articulation tract. The sound signal resulting from the articulation movements gets to his ear and in his brain appears a connection between a picture of nervous excitement in the motor area and a picture of nervous excitement from

a signal arriving aurally. The environment selects from a set of movements those which are meaningful codes in the community, by means of reinforcements. Thus, relations between the articulatory movements accepted by the community and their sound patterns are strengthened. Because of such a connection, the reception of a sound code (phoneme) which comes from another person automatically causes the corresponding pattern of excitement in the motor sphere. So speech is perceived directly in articulatory terms. It may be said that speech recognition at the phoneme level is carried out by means of imitation of the articulatory movements of the other person which were excited by the accepted speech signal. The idea of considering the problem of recognition as imitation of the automata that generate the objects of recognition belongs to M. M. Bongard who stated it 20 years ago [1]. Let us suppose that there is an automata (automata is a general notion for computer) in which we put a set of images. The computer put on each image one of three labels: for example, "man", "woman". "trash", i.e. the pictures which do not belong to either of classes 1 or 2. There are other automata which need to be trained to classify the same pictures similarly. For this purpose, training is provided, i.e. presenting to automata a limited number of pictures from each class (with the indication of a class - "1", "2", or "garbage"). In other words, it is necessary to transfer the second automata to a state which imitates the automata which generated the classification of pictures, generating this recognition problem. The idea of perception of speech in articulation terms isn't something new and has a long history (it is possible to refer even to Francis Bacon).

HANDWRITING PERCEPTION

A similar approach has been used to tackle the problem of recognition of hand-written text. It appeared that the problem of recognition of hand-written letters can be successfully solved if the text is considered as a trace of the pen movement [2]. It appeared that various letters are combinations of a small number of elementary movements and that people looking at the motionless image (text) can restore the pen movement on visible and invisible parts of the trajectory. The being of each letter is not a geometrical picture, but a complex of movements, and in essence recognition of hand-written letters consists of restoration of the complex of movements which led to the emergence of the visible image. Such representation is in accordance with many known facts from the area of hand-written text perception. In any written text it is easy to restore the trajectory of writing, including sections in which the pen didn't touch the paper. In aphasias, occurring due to brain damage, misrecognition of numbers and letters is sometimes observed. However, the patient can identify a letter if following its contour. Everyone easily learns the letter traced on his back with a finger (at the same time it is impossible to recognize a letter "P" made from wood and applied it to the human's back). However, the transition from the description of a grapheme as a mosaic of white and black points to its description in the form of a one-parametrical curve (the pen's trajectory) doesn't yet solve a problem. It is also necessary to choose how to describe the trajectory. For elements of trajectories of the Latin hand-written script the following eight signs are suggested : –  – . To each lower-case letter of the Latin script corresponds code – a sequence of the above-listed elements ("a" corresponds to code , "b" to code , and "g" to code ). Most letters possess different codes. As the speed of writing increases, the trajectory of letters becomes distorted, and consequently the codes change. However, these distortions are restricted by regularities. It is shown that one element can be transformed only into the next element of

the sequence presented above, i.e. by distortions of a trajectory the λ element can be transformed into either the O -element or the Λ -element; the \cup -element can change either into the Y -element or into the $-$ element **only**, etc. Thus a set of elements of a trajectory of hand-written letters (the alphabet) is established. And taking into account that the first and last elements of this row coincide, the chain of basic elements can be rolled up into a ring, i.e. elements ordered on a circle.

From this point of view, it becomes clear that the problem of recognition of printing letters has to be solved from other positions – as a problem of recognition of pictures – and is essentially something other than a problem of recognition of the hand-written text.

DRAWINGS PERCEPTION

We will consider one more problem of perception: the recognition of polyhedrons in schematic drawings. The drawing consists of straight lines representing visible edges of a polyhedron. The task is to determine what polyhedron is represented in the image. For finding an adequate language of description it is necessary to describe the process which creates the object of recognition, that is the polyhedrons. Such an approach looks, at first sight, quite unnatural. The situation significantly differs from the two other tasks described above, in which the generators of objects of recognition were people (to be exact, a hand of the person with a pen or its articulatory device). But let us trust in the theory. We will begin with a simple case – a prism. One way to describe what a prism is goes as follows. Let us consider a flat polygon. By stretching this polygon along a straight line (not lying in the polygon plane) a prism will be created. Lateral edges are traces of movement of corners of the base of the prism, but on the schematic image only visible edges are represented. Thus, we come to the following description of the image of a prism: a set of parallel lines (edges) leaning on the base. Such a description is constructive: it is a clear instruction for a program that can recognize a prism in a given schematic image. Now the classification of prisms is reduced to classifying the basic polygon. Thus, for a particular problem of recognizing prisms we found the description we were seeking (for example, "a prism with a triangular base" is also the name of a class to which it belongs). This approach can be used more generally for the wider class of polyhedrons [3]. First, for the truncated prisms lateral sides on the image do not necessarily have to be parallelograms – they can be quadrangles. Second, for recognition of pyramids it is necessary to search not for a system of parallel edges, but for a system of the edges proceeding from one point (for truncated pyramids the continuation of edges have to be crossed in one point).

Considering communication by means of speech, writing, and drawing suggests an idea that **in other areas of communication the perception of information occurs through the language of physical activity by internal imitation**. As an example we will point to the perception of sports. It is known that, when watching soccer or boxing competitions, the viewer involuntarily imitates a kick or a punch, i.e. empathy goes into motor language. The perception of ballet is probably based on internal imitation of observed movements as well. From the same point of view it is possible to consider the perception of music too. One violinist told me that when he listens to music, he bends and shakes his body and moves his hands and fingers. Certainly, these movements may not be shown externally, or they may be shown to a minimal degree, i.e. occurring at the level of an internal innervation. In Delgado's book "A brain and a conciseness" [4] an example is given of one person who after

an operation on his nervous system said that he now perceives music with one half of his body only. In some way this phenomenon is similar to the known fact that children writing from dictation make more mistakes if they clamp a pencil in their teeth, i.e. when participation of the articulatory organs in perception of speech has obstacles. In general the perception of music through movement is very natural. Dancing rhythms are especially easily perceived by listeners who are not expert in music . What a pleasure it is not only to hear music, but at the same time to see its performance, to witness a surprising harmony of movement and sound. On the contrary, we are jarred by the divergence between sounding speech and visible articulation which we sometimes see in the movies.

It is possible that, by teaching a child to play the piano, we not only enrich him in terms of knowledge of music and acquaintance with musical culture, but we also expand the language of his perception of music, in particular, because he becomes the owner of a much bigger set of movements, and gains a richer perception of music through the motor sphere. The same applies to perception of sports or ballet. The athlete or the dancer perceives far more nuances in observed movements than the ordinary viewer because they possess a far greater alphabet of movements. In general, it should be noted that M. Sechenov also put forward the idea that the perception of the outside world is formed to a large degree by means of muscular feeling ("dark muscular feeling"). He spoke about perception of space (distances) through the muscles of the eyes and about perception of time through movements (walking). Facts which were mentioned above say that muscular feeling plays an important role in the sphere of communication between people as well. We will now try to consider from these positions a question of communication in fauna. If communication language in a substantial measure is based on movements, it is clear that an important role is played by a physical form of existence of an individual in a choice of language of communication. But then we have to realize that animals, possessing a motor complex different from the human, can have also a different language of communication and, if we accept Sechenov's idea, other presentations of the world. Therefore it seems inadequate when many scientists (first, Lorentz) try to describe reason of animals directly in the same terms as reason of people, and try to describe their behavior – and particularly their relationships – in the same terms as for people.

MAN-ANIMAL COMMUNICATION

In the same way mutual understanding between man and animal becomes complicated because of dissimilarity of their physical organization, and Homo sapiens, professing to possess the highest intelligence in fauna, would have to show initiative and try to understand what language is most acceptable for carrying on a dialogue with other living beings. After all we still talk to dogs in human language. We admire that the dog distinguishes some dozens of words and expressions. It has to be true that the dog doesn't perceive the words in a phonetic code (not least because he has no opportunity to imitate them). The dog can use for distinction of words such characteristics as word length, accent position, existence of fricative or explosive sounds, and intonation. It is clear that by means of these signs the dog isn't able to distinguish more than a hundred words. (By the way, in recent years some computer programs for speech recognition appeared which are based on the same principle, and sure enough they recognized no more than a hundred words. Such methods may quite reasonably be called "dog-like recognition"). It is possible that the mental capacities of a

dog allow him to distinguish far more signals but only if these signals are coded in language natural to a dog. Even more difficulties would arise if we were to try to encounter beings from other worlds – their physical appearance will not resemble ours, so the problem of finding a common language will be extremely difficult.

Let us consider one more situation. The duckling has to understand well his siblings or mother if he is himself to learn to “talk”. He interprets an arriving signal adequately, i.e. in language of movements of his own articulation device. However, it has been shown that the duckling learns the birdcall of his mother as soon as he hatches from the egg, and doesn't react to similar calls of ducks of other species. The principles of imitation do not apply here, resulting in the theory's failure. However, further research shone a different light on these facts. (Bird 2014) proved that the ability to recognize sounds in ducklings of a type of *Anas platyrhynchos* arises in an embryonic state five days prior to birth. In one or two days, still in the egg, when the embryo starts breathing, it can already make sounds. Further it became clear that if the embryo is deprived of hearing the sounds he makes, then the baby bird's ability to distinguish an invocatory voice of his mother from the voices of the ducks belonging to other species is broken (we have no more detailed description of this experiment). Thus, if the embryo has no opportunity to establish a connection between the motor act and a sound picture accompanying it, perception of further sound signals of the same type is worsened. If we believe in the idea of communication by means of imitation, we have to try to come into deeper contact with a dog. We will organize for the dog a phonetic code, which uses sounds (and articulations) which are available to dog's articulatory tract. We also have taken in consideration the convenience when using them in sequence. Than a correspondence between phonemes of our language and dog's phonemes has to be established We will use a computer that will substitute sounds of our speech by corresponding sounds of dog's alphabet, and thus produce a stream of sounds composed of dog's phonemes (i.e. sounds, which the dog is able reproduce). Earphones will be placed on puppy's head. If puppy lives permanently with us in the home, he will permanently be immersed in linguistic environment. I hope that as all mammals and birds he has an instinct of imitation - the most powerful tool for learning. I hope that as time will going he will start to repeat some passages of that speech stream. We will also put earphones on our heads, and supply the computer with a program that transforms dog's sounds in human's phonetic alphabet.

CONCLUSION

In conclusion (not as confirming arguments, but as illustrating examples) I will add some facts. Reading cursive writing becomes possible for children at the second or third year of school– when they master handwriting as a motor act (synergy). It is natural that preschool children train in reading printed letters, which are distinguished as a picture and don't demand for perception the corresponding motor skills. As a rule, at such a stage reading takes the lead over writing. It is rather interesting also to remember the report on television about the boy of four years, fluently played very many difficult things on the piano, and learned them himself. This occurs in very specific situations. The child from the earliest age spent time in a bed which was situated close to a piano. In the house, music from the radio or parents playing piano was constantly heard. Thus, there was a system in which there was 1) a potential generator of musical sounds - hand of the child plus piano, 2) an external

source of musical sounds of radio, and of parents playing piano, and 3) an ear plus brain of the child, which could compare the sounds made by him, and musical sounds of the outside world. Aspiration to imitation (which is very important also when the child learns to speak) plus encouragement of parents made possible mastering playing piano in the same natural way as mastering speech.

REFERENCES

1. Stetson, R. H. (1905). A motor theory of rhythm and discrete succession: I. *Psychological Review*, 12(4), 250-270. <https://doi.org/10.1037/h0071810>
2. Liberman, A. M. (1957). Some results of research on speech perception. *Journal of the Acoustical Society of America*, 29, 117-123. <https://doi.org/10.1121/1.1908635>
3. D'Ausilio, A., Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., & Fadiga, L. (2009). The Motor Somatotopy of Speech Perception *Current Biology*, 19 (5), 381-385
DOI: 10.1016/j.cub.2009.01.017
4. Bongard, M. M. *Problems of Recognition*. Moscow: Nauka, 1967. 320 p.
5. Guberman S. and Andreevsky E., 1996, from Language pathology to Automatic Language Recognition ... and Return. *Cybernetics and Human Knowing*, 3, 41-53.
6. Guberman, S. Gestalt theory of cognition and “solid” sciences: Physics, Mathematics, and Artificial Intelligence. *Journal of Psychiatry and Cognitive Behavior*, May 2018
7. Delgado J. *Physical control of the mind: Toward a psychocivilized society*. (1969). *World perspectives*, vol. 41
8. Seabrook-Davison M. Observation of a female mallard (*Anas platyrhynchos*) with a brood of 29 ducklings. *Notornis*, 2014, Vol. 61: 51-53