THE SENSOR-MOTOR THEORY OF CONSTANT SPEECH PERCEPTION

GEORGE LOSIK

Dept. of Experimental Phonetics
Institute of Linguistics, BSSR Academy of Sciences
Minsk, Byelorussia, USSR.

ABSTRACT

New aspects of sensory and motor theories of speech perception are discussed. The possible mechanisms of constant speech perception by grown-ups as well as mechanisms of stuttering are proposed. In addition, the role of babble-speech for ear-development is analysed. Finally, the results of experimental verification and computer testing of a new sensor-motor theory are reported.

There exists a number of theories of speech perception. The well-known of them are sensory (G.Fant, M.Halle, R.Jakobson) and motor (A.Liberman, F.Cooper, L.Chistovich) theories. But the interest to the verification of these theories is dicreasing recently.

We suppose that there exists a false opinion that the speech signal informativity in these models was taken into account (N.Zagoruiko). Here we offer a new sensor-motor theory of speech perception which explains the mechanism of speech perception by grown-ups and by children. It also explains the formation of samples perception of speech units by a child during the latent interaction of audio-speech and motor systems.

Sensor-motor theory differs from other theories known recently according to the following theses.

Thesis 1. The invariant characteristics of speech units are not present in a common acoustic signal. They appear only on the first stage of speech processing after the signal normalization in audio-speech system.

Nevertheless a new informational characteristic is present in this acoustic signal, making the speech recognition invariant. This characteristic reflects the speech unit variations. A man analyses and memorizes these characteristics at an early age. Thanks to this new information a grown-up reconstructs an initial signal from distorted one during the recognition. Only after this process the reconstructed signal is compared with the sample. We confirm that speech units differ in manner of variation (Δy-description), but not only by their motor program (y-description). Δy-information forms an additional characteristic axis. Recognition in audio-speech system is realized with the help of Δy-description but without the y-description itself.

Thesis 2. The information about pronounciation variation of speech units is memorized in the audio-speech system, but not in a motor one according to a motor theory. Finally, the distortion model of articulation system is formed in the audio-speech system.

Thesis 3. A child can not reveal the variation rules of a speech signal listening to the speech of grown-ups. The acoustic signal is an indefinite function, which is formed by a countless multitude of a small number of argument combinations, i.e. speech commands.

Therefore speech of grown-ups is not used for learning the speech unit variation rules. Only elementary signal variations not the combinational ones are used by children for memorizing. Only directions of variations not the zone of variations (S.Dzhaparidze, I.Zimnyaya) are informational for memorizing.

Hence a child must reveal these rules indirectly. To learn the articulation distortions in the speech of grown-ups he imitates these distortions in his own motor-speech system.

Thesis 4. The repetitions of sounds in babble-speech (e.g. va-va-va, ba-ba-ba, ma-ma, pa-pa) are the acoustic signals, which reflects child's imitation of articulation distortions. These repetitions in babble-speech (i.e. iterations) are initiated by child's audio-speech system, but are accomplished in the motor system according to its degrees of freedom. In the state of motor system jumps are initiated. We think the iteration of the type "va-va" reflects such jump on the acoustic level. Child's auditory system perceives and analyses the multitude of iterations in his own babble. Finally this system accumulates the information about the direction of all elementary jumps. In this process a pair of acoustic signals, not a single one, becomes informative.

Thesis 5. The information about articulation distortions is mastered by a child in his early childhood only, but later it is used by him constantly. During the recognition this information enables him in the sensory system to modulate the multitude of motor work divergences and to realize the selective fitting of the input signal to the sample. During this fitting the statistical information about signal variations enables a child to discover the most probable track in which realizing distortion took place. The degree of additional information depends on the probability of track-fitting of input signal to sample, but not on the fact of their fitting. We confirm, that after this sensory learning the motor system is not necessary for recognition.

According to the sensor-motor theory a child forms sound image of a speech unit in his audio-speech system in three stages. During the first stage a child perceives the pronunciation of a word or a syllable articulated by grown-ups. He listens to many separate realizations of a speech unit and forms its average sound sample. In the speech of grown-ups the child hears distorted to a small degree realizations of a speech unit. Therefore they are grouped with little dispersion about an average value. During the second stage a child adopts the skill of pronunciation of speech units, the samples of which were formed in the auditory system. As a result, in the speech-motor system the motor samples of speech units are formed. The thesis about the existence of the third stage is new. At this stage the sensory system receives information about the modifications of those samples, which were formed in it. A child listens to the iterations of his own babble-speech. The development of iteration mechanism in the babble period consists of several substages. At the first one the child exercises the articulation of samples. Here the development of babble is going through without ear participation. Therefore this substage is presented in the babble-speech of deaf children. During this period the inborn program of articulation exercises under cinesthetic control is realized (V.Beltyukov). Therefore at first the child's babbles are realized without iterations in any acoustic situation and not only in stillness. At the next substage the child's ear begins to control the sounds of his own babble. The auditory system adopts a single babble syllable which is pronounced at this moment. Next moment the system stimulates the repeated pronunciation of this syllable. This substage is absent in the babble-speech of deaf children. On the contrary, by the children with ear perception the phenomenon of autoecholaly is developed. This phenomenon is realized in stillness mainly. The babble sounds are combined into iterational chains (va-va-va). However the autonomous mechanism of program repeated triggering is gradually formed in the motor system. After that this mechanism generates the iterations itself without ear participation.

The variation imitation in the motor system becomes possible just at this period. The direction of possible sample variations starts to be coded in double iterations. To our mind, the iterations' mechanism, that is formed with child during his babble, is kept then with grown-up in the blocked state for the whole life. This mechanism can manifest itself again. One of these manifestations is the clonical form of stuttering and Lee effect. Besides the iterations' mechanism promotes indirectly that word-iterations (mamma, papa) take the main place among the first child's words in many languages.

There are some confirmations of sensor-motor theory. One of them is the fact of infringement of ear development by a child caused by the blockade of the babble stage during the speech development. The blockade of listening to iteration in this period is manifested in the underdevelopment of the phonematic ear of child (V.Beltyukov).

We have conducted some special research of child's babble-speech, the results of which prove sensor-motor theory. Firstly, the results received, that babble iterations are realized during stillness mainly. Secondly, they reveal that double iteration of a syllable is most widely spread. Thirdly, it is shown that sound iterations diads are double triggers of one and the same motor-program. And at last, our results illustrate that iterations in the form of diade exist in the babble of children speaking different languages.

On the basis of sensor-motor theory a new explanation of babble function in speech ontogenesis can be given. The theory explains also sound ontogenesis dissociation in the child from the point of view not only of his speech development (V.Beltyukov), but of his phonetic ear too.

A corresponding mathematical model has been built for the presented theory. This model allows to test new algorithms of teaching and speech recognition in computer experiments on natural speech, to use them in the existing automatic speech recognition system.