

SPEECH PERCEPTION: OVERVIEW OF THE MODELS

Inna Krasovska

Taras Shevchenko National University of Kyiv, Ukraine

3488778@gmail.com

The paper outlines the approaches to the study of speech perception resting special attention on cognitive framework. Speech perception is analyzed through the intonation clues that ensure its multi-layer realization during speech communication. The work singles out and describes the most authoritative models of speech perception within cognitive paradigm of phonetic research.

Key words: *perception, speech, intonation clues, cognitive operations.*

Perception of speech is a complex process that can be researched through the insight into its cognitive mechanisms. This approach is ensured by the synergy of experimental data and methods from various sciences: physiology, cognitive science, linguistics and psychology. Perception is physiologically represented by a complex functional system, which is based on a dynamic sequence of links located at different levels of the nervous system. Meaningful perception, being a mental function of this physiological substrate, is also characterized by a level structure – both the stepwise nature of the process itself and the leveled hierarchy of speech signal processing (Vygotsky, 1962).

In the framework of cognitive paradigm perception is, first of all, a physiological process that consists of mechanisms that, on the one hand, ensure the perception of speech, and, on the other hand, its generation. So, this analytical-synthetic process consists of two stages that enable the synthesis and analysis of sound information. At the first stage, auditory analyzers are activated, which isolate certain characteristic features in the sound message and combine them into complex deep representations. At the second stage, these deep structures are analyzed thanks to cognitive operations and correlated with certain contents, which are elements of the speaker's world picture (Krasovska, 2009, p.15-20). Thus, the key objective of cognitive approach is to research into mental mechanisms that ensure realization of this multi-layer process and find out linguistic units which code and decode information flow.

As a result of multiple “distinctive” operations, an acoustic information is decoded in the mind and the meaning is extracted relying on «perceptual» clues formed on segmental and suprasegmental layers of the language. It should be noted that this approach resonates with F. de Saussure’s statement that “only the acoustic impression has value” (Saussure, 2011, p.11). According to the scientist, only the synergy of the methods of various sciences makes it possible to penetrate into the

essence of the material phenomena of speech. This idea is confirmed by the modern turn of linguistic research towards experiments involving the methods of psycholinguistics and cognitive psycholinguistics, which have changed the scientists' perception of the elements of speech and their interaction during the generation and perception of a message.

Analysis of perception is impossible without acknowledgement that emotion is a part of the linguistic context, reflected in behavioral traits, mannerisms, attitude towards others, facial expression, and in language expression, intonation in particular. At the same time, the study of the communicative functions of intonation proves that suprasegmental units mark a significant number of non-propositional content categories in speech, namely: logical-communicative blocks of statements (rhema/theme, mode/dictum, juxtaposition/contrast); modality, connotation, rhetorical instructions, expressive and iconic meanings. Perception of acoustic information is approached differently. There are various models of perception that rely on a comparatively different range of principles. However, we believe that they can be classified into three major groups, focusing on (Krasovska, 2023):

- cognitive operations involved in the process of perception and generation of an utterance (Chomsky, 2002);
- type of information enabling a relevant coding and decoding of acoustic information (Laver, 1994);
- contextual parameters defining the process of perception (Chafe, 1994).

In the framework of the first approach the sound pattern of an utterance perceived by speaker is considered as a part of a complete formal model that involves transformation of the surface structure of a sentence into its deep structure via cognitive operations aimed at deducting the meaning. Thus, the speaker can understand the sense of structurally different sentences, such as:

- (1) *Mr. Smith ordered the coffee.*
- (2) *The coffee was ordered by Mr. Smith.*
- (3) *It was the coffee that Mr. Smith ordered.*

In accordance with experimental data speakers extract a meaningful information relying on mental operations based on such an algorithm: $A \rightarrow B / X \rightarrow Y$, where A identifies the unit subject to the rule; B indicates what form this unit should take; X and Y denote the contexts for A, located on the left and right. Generally, the rules are formulated in terms of feature matrices, not segments, and therefore affect any class element characterized by that feature (Chomsky, 2002, p. 30-45).

A somewhat different perspective is applied to research into perception within the approach developed by W. Chafe. Thus, consciousness is believed to focus on

some important stretches of information in a given context that activates mental operation of its perception and decoding. Acoustic information is sifted by consciousness as a three-dimension body that contains such plans as: new-active-given. These information plans are reflected differently in the structure of the language. For example, given spans of information are realized in the form of weakly stressed pronouns and functional words, while active ones are strongly accented notional words. From this viewpoint, oral discourse is generated not as a smooth stream, but in spurts, quanta. These quanta, most often commensurate in volume with one predication, are called intonation units (IU). Each IU reflects the current focus of consciousness, and pauses or other prosodic boundaries between IUs correspond to transitions of the speaker's consciousness from one focus to another. The average length of IU for the English language consists of four stressed words or syllables. A prototypical IU coinciding with the clause thus verbalizes an event or state (Chafe, 1994).

The third approach is grounded on the assumption that speech perception embraces insight into audible elements of language, establish their relationship and formulate an idea of their meaning. Perception, therefore, unfolds on two levels – perception and comprehension. In its turn, comprehension involves decoding the general meaning that lies behind the directly perceived speech flow: the process of transforming what is perceived into the underlying meaning (Cohen, 1967, p. 177-192).

For instance, the meaning of the utterances *Cold! Hot!* can be perceived differently in compliance with a situational context. It can be interpreted as a piece of advice to get dressed warmly will dress warmer, a request to close a window or a guidance in a game. And in all cases, these utterances are a predicate to reality, to different situations. Following this logic, in the course of perception the recipient establishes semantic connections between words that make up the semantic content of the given utterance. As a result of comprehension, the listener may come to an understanding or misunderstanding of the semantic content of the utterance. Understanding is defined as a two-step process, where the first step is to comprehend a general idea of an utterance whereas its second one – decode a communicative intention of an utterance (Cohen, 1967, p. 177-192).

Speaking about the complexity of the cognitive system and its role in perceived and generated speech, it is vital to stress that the expression of the speaker's communicative action acquires a certain meaning thanks to prosodic means. An experimental study of the speech of people with motor aphasia showed that the listener clearly perceives the general auditory tone of speech, associates and evaluates what he has heard through commands localized in the right hemisphere of

the brain, activated by pauses, rising and falling of the main tone, etc. So, we believe that intonation can be viewed as a structural component of speech generation as it is a complex unity of speech melody, sentence stress, tempo, rhythm and voice timbre which enables the speaker to express his thoughts, emotions and attitudes towards the contents of the utterance. These are perceptible qualities of intonation.

References

1. Chafe, W. (1994). *Discourse, consciousness and time*. Chicago: Chicago University Press. 311 p.
2. Chomsky, N. (2002). *The Sound Pattern of English*. Cambridge (Mass.): The MIT Press. 470 p.
3. Cohen, A. & Hart, J. (1967.) On the anatomy of intonation. *Lingua*. No. 19. P. 177-192.
4. Krasovska, I. V. (2009). *Lingvocognityvni osoblyvosti prosodychnogo oformlenya angliyskikh vysloven na poznachenya pozytyvnyh emotsiy (eksperimentalno-fonetychne doslidzhenya)* [Linguistic and cognitive peculiarities of intonation contours of the English utterances defining positive emotions (experimental phonetic reseach)]: dis. ... kand. filol. nauk: 10.02.04. K., 2009. 263 p. [in Ukrainian].
5. Krasovska, I. V. (2023). Intonation clues in speech perception: experimental phonetic research into English discourse. *Aktualni pytannia humanitarnykh nauk: mizhvuzivskyi zbirnyk Drohobyskoho derzh. ped un-tu im. Ivana Franka*, №69(2). Drohobych, P. 120-126.
7. Saussure, Ferdinand de. (2011). *Course in general linguistics*. NY: Columbia University Press, 260 p.
8. Vygotsky, L. (1962). *Thought and language*. Cambridge MA: MIT Press. 168 p.