The Multimodal Model of Language

The basic and primary use of language is in a face-to-face conversation, "all others are being best described in terms of their manner of deviation from that base" (Fillmore, 1981; in Clark, 1996, p. 8). The understanding of language must therefore come from research of natural language between humans speaking, or signing, face-to-face. In such situations, participants use various multimodal cues, e.g., sound, linguistic and paralinguistic features, bodily language, including facial expressions, hand gestures, and empathetic wince to convey a specific meaning and reach a common ground. Other factors, such as situational context and cultural reference, affect the use of stylistic features, vocabulary, and grammatical rules. Language is thus used in a multimodal manner, and from the perspective of multimodal communication, psycholinguistics and cognitive sciences it is also considered a multimodal phenomenon (O'Connell et al., 1990; Lakoff and Johnson, 1999; Barsalou et al., 2003; Vigliocco et al., 2014; Skipper, 2014; Jacobsen, 2015).

Today, in the field of language learning, applied linguistics, sign language studies, and artificial intelligence, the traditional language models as suggested by De Saussure and Bühler are insufficient. They describe language from the perspective of a speaker and a listener, without considering all other modes of communication that are naturally part of human language. For instance, De Saussure (2011) considers language as a linguistic sign that unites a concept and a sound-image, whereas the sound-image is not a material sound but a psychological imprint of the sound and the impression that it makes on our senses (p. 66). Bühler (2011) views language as organum, i.e. "for the one to inform the other of something about the things" (p. 30). According to this, language is a production of acoustic phenomena consisting of three largely independently variable semantic relations: expression, appeal and representation. This article proposes the Multimodal Model of Language (see Figure 1). This model views language as a multimodal phenomenon, consisting of meaningful spoken or signed utterances, i.e. not of words in isolation, but of words representing a meaningful communicative act affected by multiple cues and senses present in the production and perception of language.
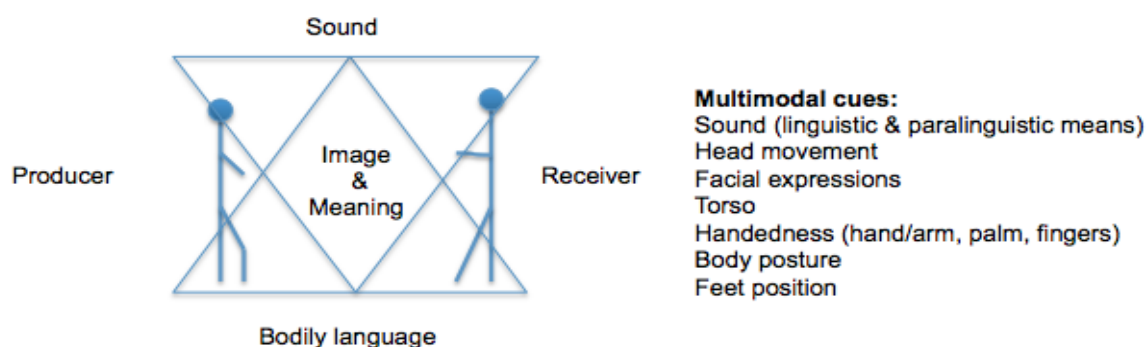

Figure 1: Multimodal Model of Language

In order to design such a model, the Complexity Theory (Filipović, 2015) together with Jörg's (2011) three complexities of human conversation were considered: (1) conveying information, (2) understanding information, and (3) building a common ground. The results indicated that the exchange of messages in a face-to-face conversation between participants does not only depend on the spoken word, what it symbolizes and how it appeals to the listener, but also on the situation of time and place, emotional and social status of the speakers, their individual knowledge of the world, and the context in which language is produced. The suggested model considers both participants

sharing information in a multimodal manner; both construct meaning by employing various senses (modes) and references to the general and specific context, in which language happens.

References

Barsalou, L. W., Simmons, W. K., Barbey, A. K., Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. In *Trends in Cognitive Sciences* 7(2), pp. 84-91. Elsevier.

Bühler, K. (2011). *Theory of Language. The representational function of language.* Goodwin, D. F. and Eschbach, A. (Translators). John Benjamins Publishing Company.

Filipović, J. (2015). *Transdisciplinary Approach to Language Study: The Complexity Theory Perspective*. First Edition. Palgrave: Macmillan

Fillmore, C. J. (1981). In Clark, H. H. (1996). *Using Language*. Cambridge University Press.

Jacobsen, M. H. (2015). Goffman's Sociology of Everyday Life Interaction. In Jacobsen, M. H. and Kristiansen, S. (Eds.), *The Social Thought of Erving Goffman* (pp. 67-84).

Jörg, T. (2011). *New Thinking in Complexity for the Social Sciences and Humanities: A Generative Transdisciplinary Approach*. Springer. Netherlands.

Lakoff, G. and Johnson, M. (1999). Philosophy in the Flesh: The Embodied Mind and Its Challenges to Western Thought. Basic Books, New York.

O'Connell, D. C., Kowal, S. and Kaltenbacher, E. (1990). Turn-Taking: A Critical Analysis of the Research Tradition. *Journal of Psycholinguistic Research* 19(6).

Saussure, F. de. (2011). Course in General Linguistics. Baskin, W. (Translator). Meisel, P. and Saussy, H. (Eds.). Columbia University Press.

Skipper, J. I. (2014). Echoes of the spoken past: how auditory cortex hears context during speech perception. In *Philosophical Transactions of the Royal Society* B 369, pp. 1-19. Royal Society Publishing.

Vigliocco, G., Perniss, P., Vinson, D. (2014). Language as a multimodal phenomenon: implications of language learning, processing and evolution. In *Philosophical Transaction of the Royal Society* B Biological Sciences 369(1651), 20130292, pp. 1-7.

492 Words