On the corpus-based method of investigating the semantics of emphatic particles (on the material of the Besermyan dialect of the Udmurt language)

This paper looks at the adaptability of different methods of studying semantics of emphatic particles in the Besermyan dialect of the Udmurt language (Uralic language family).

Two of the most prevalent methods in linguistic fieldwork are elicitation and corpus analyses. Due to the ability of the particles to have vague meaning, which can vary greatly in different contexts, it can be complicated to ask native speakers direct questions as regards the semantics of the particles . Two of the following characteristic features of discourse markers (DMs) – since in a broad sense, emphatic particles can be regarded as discourse markers – make it appropriate to use corpus analysis for investigating semantics of the units:

1. DMs "are a feature of oral rather than written discourse and are associated with informality" (Jucker & Ziv, 1998).

2. DMs "appear with high frequency" (Jucker & Ziv, 1998).

In our research, the Corpus of the Besermyan dialect of the Udmurt language was used (Besermyan corpus, version of 2012, approximately 40 000 words, compilers O. Biryuk, R. Idrisov). It contains spoken monologic and dialogic texts recorded from Besermyan-speaking people (Shamardan, Udmurtia, Russia), translated and glossed in Russian.

An analysis of the corpus data can reveal some syntactic features of a particle providing us with the relevant information about the ability of a particle to be part of a syntactic constituent (for clitic particles) or form a new DM, when used together with another particle.

Corpus data allows us to make a questionnaire based on corpus examples, when the native speakers are asked to give a context or comment on a sentence, which contains a particle.

When the core meaning of a particle is defined, it makes sense to construct a questionnaire based on the research of DM's with similar meaning in other languages. Thus, for Besermyan particle *uk* semantic descriptions of the corresponding Russian *же*, *ведь*, *а* were applied.

As the notion of a DM's scope is relevant to studying the semantics of a DM, an experiment with change of a DM's position in a sentence can show interesting results.

(1)	So	ber-e	so	proval-ti=ik	ber-lan'	tə̑b-i-z.
	that	back-ILL	that	passage-PROL=EMPH	back-APPROX	walk.up-PST-3SG
	'Then he walked back up the same side (of the) street.'					
(2)	So	ber-e	so=ik	proval-ti	ber-lan'	tə̑b-iz.
	that	back-ILL	that=EMPH	passage-PROL	back-APPROX	walk.up-PST-3SG

'Then he (the same man) walked back up the side street.'

As the result of analyses, all the corpus contexts should be assigned with one of the detected meanings. On the other hand, not all particular meanings, confirmed by elicited examples, find correspondences in the corpus. That is why we need to apply additional methods of elicitation to get the fuller view on the semantics of a particle.

Jucker, A. H., & Ziv, Y. (Eds.). (1998). *Discourse markers: descriptions and theory*. John Benjamins B. V.