

Cognitive Conceptualization in the Russian Diminutives Referring to ‘Wind’, *Ветерок*, ‘Rain’, *Дождик*, and ‘Snow’, *Снежок*: A Corpus Analysis Conducted Through R

The present contribution investigates the cognitive conceptualization of the Russian diminutives referring to ‘wind’ (*ветерок*), ‘rain’ (*дождик*), and ‘snow’ (*снежок*), adopting a cognitive linguistic approach primarily grounded in Langacker (1999) and Taylor (2003). According to this perspective, while diminutive morphemes are commonly assumed to indicate the small size of an entity, they also convey various other meanings, such as an affectionate attitude expressed by the speaker or, conversely, a sense of diminished value. Therefore, the morphological category of diminutives can be regarded as inherently polysemous. On this basis, a corpus analysis conducted using R examines the distribution and semantic behaviour of *ветерок*, *дождик*, and *снежок*, which belong to the same semantic field. Consequently, they allow for a controlled comparison of potentially privileged meanings, as well as salient similarities and differences in their usage.

The study began with the retrieval of 300 occurrences for each *ветерок*, *дождик*, and *снежок* from the Russian National Corpus. They were collected in reverse chronological order to ensure the inclusion of the most recent data. The subsequent R-based analysis examined the distribution of the three diminutives across medium, sphere, text type, and morphological inflection. With the exception of morphological inflection, these parameters were also analysed diachronically by year of publication. The results for each diminutive were further compared to identify similarities and differences in their situational use. The analysis revealed several noteworthy patterns, including a significantly higher frequency of the three diminutives in the literary sphere. Diachronic results were also of particular interest: for example, they showed that while *снежок* continues to occur in literary genres in recent years, *ветерок* and *дождик* appear in 2020–2021 predominantly in publicistic and scientific genres rather than in literary ones.

An initial qualitative examination of the relevant contexts for each occurrence of the diminutives suggested that these distributional discrepancies may be related to differences in diminutive meaning and function. To investigate this hypothesis, each occurrence of *ветерок*, *дождик*, and *снежок* was further annotated for its semantic interpretation and syntactic function, drawing on existing corpus annotations for the latter. Meaning labels were assigned based on contextual and situational factors, as well as the speaker’s evaluative stance. These labels encompass both context-dependent pragmatic and connotative interpretations and conventionalized lexical meanings.

The statistical distribution of meanings for each diminutive was then analysed using R. In addition, meanings were examined in relation to morphological inflection, medium, sphere, and text type to identify potential correlations between specific semantic values and contextual features. Moreover, the interaction between meaning, syntactic function, and morphological inflection was analysed for each diminutive to assess whether functional or grammatical factors influence semantic interpretation. This analysis revealed interesting correlations. For example, it has shown that the diminutive *снежок* predominantly lexicalizes a concrete referential meaning (‘snowball’), accounting for 48% of its total occurrences, and frequently functions as a syntactic subject, thereby illustrating how one specific semantic interpretation correlates with a grammatical function in context. Moreover, the analysis of the diminutive *ветерок* has shown that, when used to convey a negative semantic evaluation of a rather adverse situation, it frequently occurs in the genitive case and functions as an oblique cause.

Finally, meaning distributions were compared across the three diminutives to identify similarities and differences in recurrent interpretive patterns among functions, inflection, and contexts. Overall, the analysis has shown that central argument positions and structural cases (especially nominative subjects) tend to favour figurative and evaluative interpretive values, while peripheral or attributive positions and non-structural cases favour affective, attenuative, or lexicalized interpretations.

Therefore, by combining a cognitive theoretical framework with a reproducible corpus analysis in R, the study provides an account of how diminutive meanings emerge from the interaction between morphology, syntax, and usage context.

Main References

- Dressler, W. U., & Merlini Barbaresi, L. (1994). *Morphopragmatics: diminutives and intensifiers in Italian, German, and other languages* (Reprint 2011). M. de Gruyter.
- Langacker, Ronald W. *Grammar and Conceptualization*. Mouton de Gruyter: Berlin, 1999. Print.
- R Core Team. (2025). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing.
<https://www.R-project.org/>
- Russian National Corpus, (Национальный корпус русского языка) [Web]
<https://ruscorpora.ru>
- Taylor, John R. *Linguistic Categorization*. Oxford: Oxford University Press, 2003. Print.