Where Alice fell into: Motion events in a parallel corpus
Annemarie Verkerk
Max Planck Institute for Psycholinguistics

One of the fundamental elements of human experience that can be expressed in language is motion. Motion events like ‘diving into the sea’ or ‘running out of the forest’ can be expressed in every language. Several linguists have noted that there are alternative ways to encode different semantic aspects of the motion event. In Spanish, for instance, the path of movement is typically indicated by the verb, while in English, manner of movement is typically expressed on the verb while path is expressed with particles. An example is presented below.

*English original (Through the Looking-Glass and what Alice found there, by Lewis Carroll):*
So they shook hands, and then the Knight rode slowly away into the forest.

*Portuguese translation:*
Apertaram então a-s mão-s, e o Cavaleiro afastou-se, cavalgando lentamente pel-a floresta.

We are approaching the linguistic variation that exists in this domain of inquiry from a historical perspective. Within this perspective, our primary aim is to find out to which extent this variation is determined by the phylogenetic relations of each language. Investigation of patterns of common descent and diversification within the language family allow us to assess both the influence of vertical transmission on this domain and the influence of typological universals. The use of phylogenetic comparative methods allows us to model how the encoding of motion events has evolved through time and space by mapping structural diversity onto a phylogenetic tree. Phylogenetic inference tells us how this diversity has come about and which changes in the encoding have taken place. These methods also allow us to test hypotheses about the ancestral state of the encoding of motion and the reasons behind major changes in the encoding.

In order to perform this task, we are in need of a dataset that 1) allows us to establish some measure of comparison between the languages, and 2) allows for a quantitative, fine-grained perspective on motion events. Parallel corpora are highly suitable for this task. The materials on which this study is based consist of a parallel corpus of translations of three literary works: *Alice’s adventures in Wonderland, Through the Looking-Glass and what Alice found there* (both by Lewis Carroll) and *O Alquimista* (by Paulo Coelho). We have selected 300 sentences from these novels for our analysis. Currently, there are translations into 10 Indo-European languages in our sample, to which new Indo-European languages will be added in the future.

In this paper, I will present some preliminary results from our research into the evolution of the encoding of motion events in Indo-European. In parallel with the most recent publications in this field of study, we are looking at both between-language and within-language variation. This allows us to have a fine-grained perspective on the encoding of motion events. Neither the diversity displayed by the encoding of motion events nor the mechanisms by which this has come about have been analyzed from a comparative language-family perspective before. Ultimately, understanding these processes of change contributes to the understanding of the mechanisms behind typological diversity and the way in which our cognitive and conceptual domains are structured.