Evolutionary change in motion event encoding in Indo-European
Annemarie Verkerk
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
International Max Planck Research School for Language Sciences, Radboud University Nijmegen, The Netherlands

Motion Events
This poster presents some initial findings of my research into linguistic change in motion events in IE. My starting point is Talmy (1985, 1991), who postulated the difference between satellite-framed and verb-framed languages. Satellite-framed languages typically express manner on verbs and path on a separate element (satellite), while verb-framed languages express path on verbs and manner on a separate element. An example:

English: She (Alice) ran out of the house

Portuguese (verb-framed):
saiu

Dutch (satellite-framed):
rende ze het huis uit

Based on recent proposals (Slobin 1994, Croft et al. 2010, Beavers et al. 2010), I use a quantitative approach that takes into account all the different strategies that are encountered. An overview is presented below:

Alice laughed so much at this, that she had to run back into the wood for fear of their hearing her; she had to enter the wood running path-only: she had to enter the wood manner-only: she had to run in the wood M+P verb: she had to run+enter the wood coordination: she had to run and go back into the wood other: she was in the wood

Questions
How did motion event encoding in Indo-European evolve?
1. Is there a signal of history in these features?
2. How did Proto-Indo-European encode motion events?

Data and sample
Data is taken from a parallel corpus of three novels: Alice’s adventures in Wonderland (Lewis Carroll) Through the Looking-Glass and what Alice found there (Lewis Carroll) O Alquimista (Paulo Coelho)

Languages in the current sample include: English, Dutch, German, Portuguese, French, Irish, Breton, Lithuanian, Russian, Polish, Persian, Hindi, Armenian, Albanian, and Greek

Methods
Phylogenetic comparative methods are used in this study to analyze how evolution took place. These methods employ the historical information represented by a set of phylogenetic trees to make inferences about the changes that led to the current distribution of the data. The phylogenetic trees that were used in this study were taken from Dunn et al. (2011). They are summarized below in a maximum credibility tree. The motion data was aggregated by means of a principal component analysis. The score of each language on the first principle component was taken as a score on a scale from absolute satellite-framed-ness to absolute verb-framed-ness. These scores are displayed after the name of each language.

Results
Overview of encoding strategies

1. Historical signal
Closely related languages (like biological species) tend to be more similar than languages which are not so closely related. This is called historical signal. Two methods were used:

- Pagel (1999): the likelihood of the real trees to give rise to the data is significantly different from the likelihood of trees with zero lambda (without branching structure) to give rise to the data (p < 0.01)
- Blomberg et al. (2005): the real trees are a better fit of the data as compared with altered trees in which the data have been randomly permuted across the tips of the tree (kappa = 1.42)

Both analyses show that historical signal is present and thus that motion event encoding is partly determined by history.

2. Ancestral state estimation
The hypothesis that Proto-Indo-European was satellite-framed (Acedo Matellán and Mateu 2008, Talmy 2007) was tested with a maximum likelihood method for estimating ancestral states (Pagel 1994). First, only contemporary data were used. The red bar in the figure below indicates the range of values estimated for PIE. The scores for Persian and Russian represent the two most radical scores on the scale of satellite-vs-verb-framed-ness.

With only contemporary language data, PIE was estimated to be similar to Greek and Breton. Its position on the scale from is in the middle, suggesting a language which uses a mixed set of strategies. However, since Indo-European is a family with a well-attested history, it is possible for us to add information about motion encoding strategies in ancient languages.

Sources on Latin (Watts 1964, Vincent 1999, Iacobini & Masini 2006) and Sanskrit (Watts 1964, Bloch 1963) point out that these languages probably had a satellite-framed encoding system. We added information on these two languages (based on Latin translations of the two Carroll books) to the analysis. The results are below.

The red bar still indicates the range of values estimated using only data from contemporary languages. The blue bar indicates the range of values that were estimated for PIE taking into account the data on Latin and Sanskrit as well. The result is that the range of values becomes broader, taking into account conflicting signal from different IE branches. It also shifts to the right, indicating that it has become more likely that PIE used more satellite-framed strategies.

Future work
In future work, I will try to implement more information on ancient languages to come to a synthesis between syntactic reconstruction and phylogenetic methods of ancestral state estimation. In addition, I will perform analyses to answer questions about correlated change of motion encoding strategy and verb class size.