

Summary Thesis Edwin Wildeboer

An experiment concerning visualizing non-subordination and multidominance in syntactical tree structures.

Tree structures are a commonly used tool for visualizing hierarchy. These are also being used for analysing constituents in sentences. In a normal way the tree structures have a form which is binary and two dimensional and are commonly seen as an easy representation of different forms of hierarchical information. However, a problem arises if specific information is seen as non-hierarchical or if there is a case of multiple non-subordinates.

A specific area of usage with hierarchically ordered data is linguistics, in particular *syntaxis*. In syntaxis, relations between different kinds of word groups in a sentence are being studied and the structure of a sentence is being described.

This research project accompanies the NWO-project 'The Syntax of Non-Subordination: Parentheses, Appositions and Grafts' (M. de Vries, H. Heringa & M. Kluck) and makes use of information visualization techniques which are applied on syntactical sentence constructions.

Because it appears that in the syntaxis a simple hierarchical approach does not seem to be adequate for Coordination and Right Node Raising, another form of visualization has therefore to be looked into. This is the reason why an experiment is set up to test five different forms of visualization of tree diagrams (including 3D trees) for the representation of non-subordination and multidominance. The main question in this test is to determine useful alternatives for visualizing these syntactical tree structures.

Two terms which need explanation are *non-subordination* and multidominance. *Non-subordination* is a collective term for constructions which are not subordinate to other constructions. *Multidominance* concerns nodes which are connected to multiple nodes which are hierarchical higher in trees. With syntactical constructions where can be assumed that there is a case of non-subordination and multidominance, visualization problems may occur. In particular if we also want to apply here 2D binary trees as a representation. In linguistics there are conceptions reflecting non-subordination to syntactical constructions. Here constituents exist which don't have a subordinate relation with other constituents; this is called *coordination*.

The research question consists of one general main question which contains a deepening in the form of several subquestions. With the experiment will be tried to give an answer to the questions which are mentioned underneath:

Research question

- Which model of visualization is most useful for the representation of syntactical constructions which contain non-subordination and/or multidominance?

The above results in the formulation of several questions which concern the possible usefulness of the five forms of visualizations of syntactical tree structures (in case of non-subordination and/or multidominance). These questions are split up into a syntactical category (*Coordination, Right Node Raising* and *Non-constituent Right Node Raising*) and a question type (*Constituency, Dominance, Coordination, Parallel* and *Parallel Dominance*). There is also investigated if a difference is perceptible between the results of the different groups of participants on the test, these are "laymen", students and experts.

To get an answer on the above mentioned research question and corresponding subquestions, an empirical research in the form of an experiment was set up with 46 participants. During this experiment five models of visualization of syntactical tree structures have been evaluated by objective and subjective judgements. To determine the usefulness of the models of visualizations, three parts have been used which should be applied by the testing of usability according to ISO, which are *effectivity*, *efficiency* and *satisfaction*. This occurs by applying the values to be measured which are applied to testing of usability. The three values to be measured for this research are incorrect answers (*error rate*), answer time and subjective appraisal by the participants.

Three different groups (with expert, basic and no knowledge of linguistics) were asked for this research to answer questions which applies to constructions of Coordination and (Non-constituent) Right Node Raising and to give their preference for each model of visualization.

On the hand of the test it can be concluded that the visualization model 2D+ is the most useful one. In contrast of the expectation it appears from the results that 3D+ is the least useful syntactical tree diagram for representing non-subordination and/or multidominance.