

## Linguistic Research in the CLARIN infrastructure

Jan Odijk

Utrecht University

The Dutch adverbs *erg*, *heel* and *zeer* ('very') have identical meanings but show differences in syntax: *zeer* and *erg* can modify verbs, adjectival and prepositional predicates, but *heel* can modify only adjectival predicates. (cf. *Hij is daar zeer/erg/heel blij mee* vs. *Hij is daar zeer/erg/\*heel in zijn nopjes mee* and *Dat verbaast ons zeer/erg/\*heel*). English *very* is like *heel* in these respects; (cf. *He is very \*in love/ amorous; it surprised us very \*(much)*).

Since the words *erg*, *heel* and *zeer* are synonyms this difference cannot be ascribed to a difference in semantics; it must be a purely syntactic difference that children somehow are able to master when acquiring (their first) language.

The question is: how do children when acquiring their (first) language get to 'know' this? Occurring examples of the relevant constructions of course contribute, but how can children ever find out that *heel* cannot modify verbs or prepositional predicates? No examples of that will occur in the language use they encounter. It is not clear whether children make errors against this rule of Dutch grammar in any stage. Even if they do, it is generally assumed that children get no or hardly any corrections of their language use (negative evidence), and even if they get it they appear to ignore it.

Several hypotheses to account for this problem are possible and have been proposed. In my contribution I want to show how the CLARIN infrastructure that is currently being designed and partially already implemented will enable linguists to investigate this problem in a completely new way and making use of a far wider range of empirical material than ever before: a way of investigating this problem that was not possible at all until recently, and that currently is possible only with great difficulty and if a linguist teams up with specialists in language technology.