

Information structure in spoken and multimodal Danish data

Patrizia Paggio

University of Copenhagen, Centre for Language Technology

University of Malta, Institute of Linguistics

paggio@hum.ku.dk patrizia.paggio@um.edu.mt

In this talk I will discuss how information structure relates to aspects of spoken Danish by presenting data from two different corpora that were both annotated with labels indicating sentence topic and focus. The first part of the talk presents an analysis of the relation between pauses and information structure categories in the Danish DanPass monologues. It is shown that i. words making up the sentence topic tend to be followed by pauses, and ii. the first pause in the focus domain often occurs before the word that distinguishes the object in focus from other similar objects in the discourse. The second part of the talk addresses the question of how gestures – in particular batonic head movements – relate to information structure, and draws on data from the Danish NOMCO corpus, a collection of video-recorded dyadic interactions in which speakers get to know each other for the first time. It is shown that there is a significantly strong correlation between focus and gestures. Based on these findings, suggestions are given for how the contribution of batonic gestures could be represented in a constraint-based grammar .