Operationalizing the function of discourse markers via sequencing constraints: The case of English so

Chris Koops & Arne Lohmann

A University of New Mexico, USA
B University of Vienna, Austria

An important challenge in the study of extra- clause constituents, specifically discourse markers (DMs), is the task of objectively determining their extra- clause (i.e. discourse-level) status. The English discourse markers and, but, and so are a case in point. These markers remain formally identical to their sentence-level counterparts, i.e. coordinating or subordinating conjunctions. As a result it may be difficult or even impossible to determine with certainty whether a speaker intended for the form to be understood as marking a discourse-level relationship or a sentence-level relationship. This indeterminacy hampers the objective study of DMs, as the burden of evidence shifts entirely to qualitative observations.

In this talk, we argue that a solution to this problem can be found if we capitalize on the well-known tendency for DMs to be used in two-part sequences, e.g. now therefore, but then, etc. In earlier work, we have found that DMs in such sequences exhibit strong ordering preferences (Koops & Lohmann 2013). Typically, two DMs' preferred order is that which is predicted from the canonical order of their sentence-level counterparts, so that, for example, DMs deriving from coordinators precede DMs deriving from subordinators. Accordingly, the sequence and so is common but the sequence so and is rare. Here we are specifically interested in the conditions underlying non-canonical ordering, for example the initial position of so in so and. We test the hypothesis that occurrence in non-canonical position is correlated with a DM's discourse-level (as opposed to sentence-level) function. In this sense, ordering constraints serve as a formal indicator of DM status.

Our analysis focuses on so in sequence with and and but. We extracted all instances of the ordering possibilities of so relative to and and so from the Fisher corpus of North American English telephone conversations (Cieri et al. 2004, 2005), i.e. both canonical and non-canonical ordering. Our results show that, as predicted, in the case of canonical ordering (e.g. and so), so typically functions as a marker of result or consequence, with scope over the following clause only. When occurring in a non-canonical order (e.g. so and), it frequently has larger scope and functions to structure discourse, as in signaling topic continuation or functioning as a turn-taking device. We interpret these results as reflecting varying degrees of decategorialization of so as DM (in the sense of Hopper 1991), with an increase in positional variability reflecting so's taking on more abstract, discourse-level functions.

References


