
Annotating tense and aspect - A cross-linguistic perspective

Mark-Matthias Zymla*¹

¹University of Konstanz – Germany

Abstract

Tense, aspect and mood (TAM) have been analyzed syntactically and semantically for a wide variety of languages, e.g. Mucha (2016), Bochnak and Klecha (2015), Cable (2015, 2013), Grønn and Von Stechow (2010), Filip and Rothstein (2006), Kusumoto (2005) among many others.

The recent literature suggests that categorization in the terms of traditional syntactic and semantic work (e.g. Partee 1984, Heim 1994, Vendler 1957, Reichenbach 1947) does not necessarily provide appropriate universal TAM categories (Tonhauser 2015, Bittner 2014). Thus, we propose a cross-linguistic semantic annotation scheme that i) provides a syntax/semantics interface for natural language processing, and ii) represents the semantics of tense and aspect at an abstract, universal level employing empirically testable meaning features. To this end, we offer a small corpus consisting of annotated sentences from different language families that are parallel in meaning. The sentences are chosen to reflect different tense/aspect configurations; they are syntactically annotated in terms of Lexical-Functional Grammar output produced by grammars in the ParGram project (Sulger et al. 2013, Butt et al. 2002).

We illustrate how our annotation scheme advances research on the syntax/semantics interface with an analysis of two types of syntax/semantic mismatches: i) same morphosyntax – different semantics, e.g. when comparing Slavic-style aspectual systems to systems of other language families; ii) different morphosyntax – same semantics, for example when comparing a language with morphological tense (e.g. English) to a morphologically tenseless language (e.g. Mandarin Chinese).

Our annotation scheme is beneficial for both well-researched languages and under-resourced languages from different language families with respect to the research of their tense and aspect systems. Furthermore, the annotation scheme is designed to allow for the visualization of semantic properties of TAM markers. We thus provide the possibility to create resources for a wide variety of applications, for example information retrieval, (machine) translation and language acquisition.

Keywords: Tense, aspect, annotation, natural language processing, Syntax/Semantics interface, Lexical functional grammar, crosslinguistic research

*Speaker