

Indo-European Perfects in Typological Perspective

Östen Dahl

In collaboration with my Stockholm colleague Bernhard Wälchli, I have been working on a corpus-based investigation of perfects and related categories in the world's languages. We are using a parallel corpus consisting of translations of the New Testament into around 1100 different languages from all parts of the world. We are relying on work by Robert Östling (presently at Helsinki University) which has made it possible to link these translations at the word level. The basic idea is one presented already in Dahl (1985): based on the assumption that linguistic items (whether grammatical or lexical) with same meaning and/or function will have similar distributions in parallel texts, it is possible to use automated methods to extract language-specific "grams" belonging to the same "cross-linguistic gram-type" in parallel texts, without even knowing anything about the languages in question in advance. However, while such methods work very well for grammatical markers with a constant shape, it is considerably trickier to identify grammatical categories that involve variation in form, in particular non-concatenative morphology (such as ablaut in Indo-European). Here, it is often necessary to use one's knowledge of the system to be able to find the grammatical categories one is looking for. In the case of perfects, it so happens that Indo-European languages tend to have perfects that are among the more difficult to identify in this way. Due also to problems with scripts, our coverage of the eastern branches of the Indo-European language is not wholly optimal. Still, in our present sample of 305 categories, 23 are Indo-European.

It has been known for some time that so-called "habeo perfects", which are the most common type in the Indo-European languages in Europe, are rarely found outside Europe. Somewhat more often one finds the type represented e.g. in many Indo-European languages (sometimes combined with "habeo perfects") that derive historically from a combination of a copula with some non-finite form of the verb ("esse perfects"). But what we see in our sample is that the historical source that is probably the most common cross-linguistically are words meaning 'already' or 'finish' (this is particularly common in south-east Asia and the Pacific). These (referred to as "iamitives") however tend to have a somewhat different semantics, in that they -- like *already* in English -- are used also with stative predicates to express a current state (as in *It is already dark*) and also tend to be grammaticalized in combinations with "natural development predicates" such as 'old' and 'close (in time)'. At the same time, they tend to converge in meaning with other perfects. Interestingly, what looks like incipient "iamitives" are showing up in several European languages (such as Spanish, Portuguese, Low German, Afrikaans, and Slavic languages) in that words meaning 'already' start expanding their domain of use and encroach on the territory of perfects.

It is not only the case that European perfects tend to have an uncommon historical origin, they also are somewhat unusual in their semantics. Some uses or readings of perfects are thus less common outside of Europe: this goes in particular for certain types of "existential" and "universal" readings. Linguists who study the semantics of perfects tend to focus on English and a few other European language; although some attention is usually paid to differences between languages, scholars are usually not aware of the full range of cross-linguistic variation. Typologically oriented research may therefore present a useful corrective to parochialism in this regard.