The Evolution of Subject Licensing in Indo-European Languages

Paul Kiparsky
Stanford University

The main types of grammatical subject licensing found in Indo-European languages develop unidirectionally in the following sequence.

1. In the earliest Rigvedic and Greek, every sentence has a thematic subject (modulo pro-drop); the subject bears nominative case, unless it is clausal. This is the earliest attested stage, probably to be reconstructed also for the proto-language. There are no impersonal constructions at this stage, either active nor passive.

2. Subjectless sentences arise. At this stage, impersonal actives and impersonal passives enter hand in hand. Indic shows the beginnings of this development in late Vedic, continuing through Classical Sanskrit. Latin and Germanic are at this stage already in their their earliest recorded stages.

3. A new functional projection emerges whose specifier can license subjects (TP or ΣP). It is diagnosed by verb position, by negation, by obligatoryness of finite verbs, and, in languages with clitics, by their shift from Wackernagel position to second position in TP. Logical subjects (e.g. dative experiencers) can move to this licensing position, where they have "subject properties" with respect to construal processes (anaphora, control). This stage is reached in Middle Indic; similar developments are found in Romance, Old English and Scandinavian.

4. From this point, two paths diverge.

   1. One group of languages develops consistent case licensing by endowing one or several oblique cases with structural (direct) case status, normally as ergative case. These cases can then license grammatical subjects by themselves irrespective of their position in the clause. Subjects so licensed may acquire so-called "behavioral" subject properties, such as becoming controllable (PRO). The result is a consistent case licensing language. Some modern Indo-Aryan languages have this system.

   2. In another group of languages, case licensing is lost, and positional licensing become obligatory. Subjects must be placed in the specifier of TP. This is typically associated with (and perhaps in part caused by) the attrition of morphological case (Icelandic being a famous exception).

I propose an analysis of the different systems in a linking-theory framework, whose leading idea is that abstract Case (and in particular the abstract case [+HR] of subjects) must be licensed by unifying with a morphosyntactic feature borne by morphological case, agreement, or position. Stage 1 is characterized by two constraints:

(1)   a. A sentence has a thematic subject.

        b. A subject has nominative case.

Independent support for (1) in Greek is that dative and genitive objects of two-place predicates become nominative subjects in passives. E.g. *epibouleúō* 'plot against' assigns dative case to its object, but the dative regularly passivizes as a nominative, which agrees with the verb:

(2) põs àn epebouleúsaimi autõi   hó ti mề kaì epebouleúthēn   hup' autoũ
    how PRT plot-AOR-OPT-1SG him-DAT, unless   also plot-AOR.PASS-1Sg by   him-GEN
    'How could I have plotted against him, unless I had been plotted against by him.'

Sentential arguments receive a Theta-role, hence abstract Case, but unlike nominal arguments they can't be marked for case. I assume that they can satisfy the case licensing requirement (1b) by an associated null expletive with nominative case, which at Stage 1 can only be associated with a clause, due to (1). At Stage 2, constraint (1a) is demoted, allowing the expletive to serve as a non-thematic nominative subject. At Stage 3, constraint (1b) is demoted, allowing subjects with non-nominative case. By the licensing theory they must then be licensed in some other way as grammatical subjects, either by verb agreement or by position. But since non-nominative arguments cannot agree with the verb, they must be licensed positionally. The trajectory terminates either with a pure case-licensing system or with a pure positional licensing system.

Although the full trajectory is to my knowledge documented only in Indo-European, parts of it are found in other languages as well. Its initial and final stages appear to be structurally rather simple and uniform, and its intermediate transitional stages are complex in comparison – an interesting characteristic of many long-term historical drifts, whose explanation remains a mystery.