Tracking Case Innovation: A Perspective from Marathi

The rich inflectional case system of Old Indo-Aryan was lost during Middle Indo-Aryan (MIA), but many of the New Indo-Aryan (NIA) languages developed new case markers. NIA case-systems are generally characterized by four properties: a) (split-)ergativity; b) differential object marking (DOM); c) homophony between case markers; d) the signaling of semantic distinctions via case markers. A formal distinction between structural and semantic case is difficult to draw as even the "core" case markers (ergative, nominative/absolutive, accusative) tend to carry semantic information. With ergatives this pertains to agentivity and volition, with accusatives this tends to involve referentiality/specificity and/or animacy (Butt and King 2003, 2004) both often in Differential Case Marking (DCM) with nominative.

In this paper we use the new Indo-Aryan language Marathi as a window on understanding the innovation of new case markers. We investigate the case marking system of Old and Middle Marathi as contrasted with MIA via a close corpus-based study (digital texts of the Dnyaneshwari (ca. 1290 CE), Lilacharitra (ca. 1286 CE) and Dāsabodha (ca. 1654 CE)). We show that the earliest documented period already contains new case markers and that objects show variation in their case marking. Based on our corpus data, we analyze the variation as being due to case markers in competition, eventually leading to a complex system that is not predicted by dominant ideas on the emergence and distribution of case.

Consider the example (1) from MIA. Neither subjects nor objects are overtly marked and agreement also does not conclusively identify subjects vs. objects (the # indicates clause boundaries).

(1) #kiṃ **tamu**   haṇ-ai   ṇa vālu **ravi#**   #kiṃ vālu **davaggi**
     QUES darkness.NOM.SG destroy-IMPF.3.SG NEG young sun.NOM.SG QUES young fire.NOM.SG
     ṇa ḍah-ai   **vaṇu#**   #kiṃ kari   dal-ai   ṇa vālu
     NEG burn-IMPF.3.SG forest.NOM.SG QUES elephant.NOM.SG shatter-IMPF.3.SG NEG young
     hari#   #kiṃ vālu ṇa ḍaīk-ai   uragamaṇu#
     lion.NOM.SG QUES young NEG bite-IMPF.3.SG snake.NOM.SG
     Does the young (rising) sun not destroy darkness? Does the young fire (spark) not burn down the forest?
     Does a young lion (cub) not shatter the elephant? Does the young snake not bite?
     (PC 2.21.6.9)

The equivalent of the third clause in modern Marathi is in (2a). A version in the past tense is shown in (2b). Marathi is a split-ergative language, with the ergative appearing when the verb has perfective morphology. The examples in (2) also show that in addition to an innovated overt ergative marker, objects can carry an overt accusative and engage in DOM according to the referentiality of the object (2b vs. 2c).

(2)   a.   siṃha   hattī=lā   mār-to
           lion.M.NOM elephant.M=ACC kill-PRES.3.M.SG
           'The lion kills the elephant.'
      b.   siṃhā=ne   hattī=lā   mār-le
           lion.M=ERG elephant.M=ACC kill-PERF.3.N.SG
           'The lion killed the elephant.'
      c.   siṃhā=ne   hattī   mār-lā
           lionM=ERG elephant.M.NOM kill-PERF.3.M.SG
           'The lion killed an elephant.'

A central question posed by these data is how a language develops a case marking system as in (2) from an ancestor stage as in (1). One possible answer that has been proposed is that languages encode case distinctions in order to unambiguously identify grammatical relations. Ideas of *markedness*, for example,

predict that new case markers will arise first in situations where it is difficult to distinguish agents/subjects from patients/objects, i.e., in *marked* situations (e.g., Malchukov and de Swart 2009, de Hoop 2009). The assumption that the core grammatical relations (subject/object) need to be distinguishable and that one of these needs to be marked in opposition to an unmarked nominative or absolutive case also lies at the heart of Dependent Case Theory, e.g. as articulated in Marantz (1991) and Baker (2015).

In the earliest stages of Marathi, we already see innovated case markers for the ergative *ne/ni* and for objects of transitive verbs: the clitics *-tẽ* and *-si/sī*. However, bare oblique object and nominative arguments also appear in transitive clauses. The oblique marking is a vestige of the original inflectional case marking system in OIA and is the most frequent object marker in transitive clauses. The case clitics *-tẽ* and *-si/sī* do show a difference in distribution: the majority (72%) of theme/patient verbs have *-tẽ*, while the majority of possessor/goal verbs (also 72%) take *si/sī*. Our results indicate a system that is changing from one in which oblique (originally dative/genitive marking) is used to mark both theme/patients and possessor goals to one in which distinct case-markers carry distinct semantic loads.

In Middle Marathi the highly frequent oblique marking is virtually lost. Possessor/goal verbs occur more frequently with *si/sī* (80%), but theme/patient verbs do as well (78%). In addition, a newly innovated case marker shows up: the eventual dative/accusative *lā* seen in (2) and in (3).

(3)   anu=ne   nishā=lā   pustak   di-la
           Anu=F.ERG Nisha=F.DAT book.NOM.N.SG give-PERF.N.SG
           'Anu gave Nisha a book.'

These data suggest a system in which the case-clitic *tẽ* has been replaced by the case clitic *si/sī* in theme/patient verbs. That is, a more dative oriented case clitic takes over the domain of the traditional accusative, thus anticipating a dative/accusative homophony found in modern Marathi and pervasively in many other NIA languages. Eventually, the case clitic *lā* replaces the *si/sī* as the dative/accusative marker in another instance of case markers in competition.

There are two further properties to note. For one, the agreement in MIA was ergative in alignment in the sense that the verb only agreed with a nominative argument. This basic agreement pattern is carried over to Marathi even though overt case clitics are innovated and the Marathi case system ends up making different distinctions than the MIA one. For example, the DOM illustrated in (2b) vs. (2c) is not attested in MIA. Given the close relationship posited between case and agreement in the dominant paradigm, one would expect the relationship between case licensing and agreement to change as the overall case system changed. However, this is not attested.

For another, both the Dependent Case and the markedness approach posit a close relationship between ergative marking and object marking. When there is an ergative, the unmarked option is supposed to be the preferred choice for objects. This predicts that innovated object markers should be less frequent in perfective and therefore ergative clauses. However, we found no difference in distribution of *tẽ* (the innovated marker for themes/patients) in perfective vs. imperfective clauses. Instead, accusative marking is robustly attested in both ergative and non-ergative clauses from the earliest period of Old Marathi.

Our own analysis builds on Butt and King (2003, 2004), who argue that the core insight towards understanding the distribution of case marking lies in a lexical semantic approach to case. Butt&King instead argue for an approach by which individual case markers carry semantic information that interacts with the overall semantics of a clause. This basic analysis allows the necessary room for individual variation and competition in the use of case markers as attested in the history of Marathi. Furthermore, the emergence and replacement of accusative/datives in Old Marathi simultaneously in perfective and imperfective clauses is consistent with these assumptions as the semantics expressed by DOM (referentiality) is independent of the role of the ergative case.